Hans-Hermann Dietrich

ZUR PROGNOSE DER EINNAHMEN AUS UMSÄTZEN MIT HILFE VON VERWEILZEITVERTEILUNGEN

1977

1976

Digitized by the Internet Archive
in 2026

https://archive.org/details/IA41547212_0031

ZUR PROGNOSE DER EINNAHMEN AUS UMSÄTZEN MIT HILFE VON VERWEILZEITVERTEILUNGEN

I n a u g u r a l - D i s s e r t a t i o n
zur Erlangung des akademischen Grades eines
Doktors der Wirtschaftswissenschaft
durch den
Fachbereich Wirtschafts- und Sozialwissenschaften
der Westfälischen Wilhelms-Universität zu Münster

vorgelegt von
Dipl.-Kfm. Hans-Hermann Dietrich
aus Bielefeld

1977

UNIV.
LEIDEN
BIBL.

Erster Berichterstatter:	Prof. Dr. J. Bleymüller
Zweiter Berichterstatter:	Prof. Dr. H. Gülicher
Dekan:	Prof. H. Hartmann Ph.D.
Tag der mündlichen Prüfung:	7. Juli 1976

Inhaltsverzeichnis

	Seite
Verzeichnis der Tabellen	V
Verzeichnis der Abbildungen	VI
Verzeichnis der ständig wiederkehrenden Symbole	VIII
1 Einführung	1
2 Einnahmeprognose mit Hilfe von Verweilzeitverteilungen	5
2.1 Vorbemerkungen	5
2.2 Begriff der Verweilzeitverteilung	7
2.3 Aufbau und Darstellung des Liquidationsspektrums	9
2.4 Arten von Verweilzeitverteilungen	13
2.5 Prognose mit dem Liquidationsspektrum	14
2.5.1 Prognose mit der Betragsverteilung	14
2.5.2 Prognose mit der Anzahlverteilung	19
2.6 Deterministische und stochastische Prognose mit der Verweilzeitverteilung	22
3 Ermittlung der Verweilzeitverteilung	31
3.1 Vorbemerkungen	31
3.2 Ermittlung der Anzahlverteilung	32
3.3 Bestimmung des Mindeststichprobenumfangs zur Ermittlung der Anzahlverteilung	35
3.4 Ermittlung der Betragsverteilung nach NEUMEYER	37
3.5 Ermittlung der Betragsverteilung durch Verhältnisschätzung	45
3.6 Bestimmung des Mindeststichprobenumfangs zur Ermittlung der Betragsverteilung durch Verhältnisschätzung	50
3.7 Modifizierte Ermittlung der Verweilzeitverteilung	57
3.7.1 Ermittlung aus einer vollständigen Stichprobe	57
3.7.2 Ermittlung aus einer unvollständigen Stichprobe	67

Seite

3.8 Aufbereitung der geschätzten Verweilzeitverteilung 72
3.9 Vergleich der Verfahren 75

4 Genauigkeit der Einnahmeprognose mit Verweilzeitverteilungen 79

4.1 Vorbemerkungen 79
4.2 Ex-ante-Betrachtung des Prognosefehlers 82
4.2.1 Modellmäßige Beschreibung des Prognosefehlers 82
4.2.2 Einflußgrößen der Prognosegenauigkeit 96
4.2.3 Intervallprognose der Einnahmen 103
4.3 Nachträgliche Beurteilung der Prognosegenauigkeit 107

5 Überwachung der Stabilität der Verweilzeitverteilung 115

5.1 Vorbemerkungen 115
5.2 Direkte Stabilitätsüberwachung 117
5.2.1 Überwachung mit Hilfe der Restforderungen 117
5.2.2 Überwachung durch Stichprobenentnahme 123
5.3 Indirekte Stabilitätsüberwachung 134

6 Einsatz der multiplen Regressionsanalyse unter den Annahmen der Prognose mittels Verweilzeitverteilungen 149

6.1 Vorbemerkungen 149
6.2 Schätzung der Regressionsparameter 151
6.3 Auswirkung der Annahmen über das Zahlungsverhalten auf die Schätzung 156
6.4 Beeinträchtigung der Schätzung durch Multikollinearität 161
6.5 Bestimmung der Anzahl der in den Ansatz aufzunehmenden Variablen 167
6.6 Zusammenfassende Beurteilung der Einsatzmöglichkeit 171

7 Schlußbemerkungen 173

Literaturverzeichnis 177

Verzeichnis der Tabellen

		Seite
Tab. 1:	Rechenschema zur Berechnung der Einnahmen	16
Tab. 2:	Beispiel für ein Schema zur Kumulierung von Rechnungsbeträgen	38
Tab. 3:	Symbole der Größen der dreistufigen Stichprobenerhebung	42
Tab. 4:	Matrix der Ergebnisse einer Stichprobe zur Ermittlung der Anzahlverteilung	58
Tab. 5:	Matrix der Ergebnisse einer Stichprobe zur Ermittlung der Betragsverteilung	61
Tab. 6:	Matrix der Ergebnisse einer Stichprobe zur Ermittlung der Anzahlverteilung, die nur aus bezahlten Rechnungen gezogen wurde	67
Tab. 7:	Beispiel für die Verteilung von Rechnungsbeträgen in der Grundgesamtheit	89
Tab. 8:	Beispiel für Verteilung von bezahlten Einzelbeträgen in der Grundgesamtheit	90
Tab. 9:	Arbeitstabelle zum Vorzeichen- und run-Test	121
Tab. 10:	Arbeitstabelle zum KOLMOGOROFF-SMIRNOV-Test	125
Tab. 11:	Tatsächliche Liquidation der Periodenumsätze	137

Verzeichnis der Abbildungen

		Seite
Abb. 1:	Phasenfolgen	5
Abb. 2:	Darstellung einer Verweilzeitverteilung als Histogramm	12
Abb. 3:	Darstellung einer Verweilzeitverteilung als Häufigkeitspolygon	12
Abb. 4:	Zusammenhang zwischen Umsätzen und Umsatzeinnahmen beim Liquidationsspektrum (s_0 s_1 s_2)	15
Abb. 5:	Grundgesamtheit bei der Ermittlung der Verweilzeitverteilung zu Prognosezwecken	31
Abb. 6:	Entnahmebeispiel für ein dreistufiges Stichprobenverfahren	41
Abb. 7:	Ungeglättete Verweilzeitverteilung	73
Abb. 8:	Geglättete Verweilzeitverteilung	73
Abb. 9:	Graphische Darstellung der relativen Häufigkeiten des Beispiels	89
Abb. 10:	Verteilung von B_i für verschiedene P_i	91
Abb. 11:	Regression von e_t auf $\hat{e}_t$	111
Abb. 12:	Beispiel für zwei zu verschiedenen Zeitpunkten ermittelte Betragsverteilungen	119
Abb. 13:	Kumulierte Verteilungen der Abb. 12	120
Abb. 14:	Beispiel für gestufte Kontrollgrenzen	143
Abb. 15:	Kontrollgrenzen der standardisierten Prognosefehler	144
Abb. 16:	Veränderung des Annahmebereichs von H_0 bei zu klein geschätzter Varianz $V(F_t)$	145
Abb. 17:	Veränderung des Annahmebereichs von H_0 bei zu groß geschätzter Varianz $V(F_t)$	146

Seite

Abb. 18: Zusammenhang von Prognosefehler, Verzerrung und Fehlervarianz 147

Abb. 19: Zeitreihen, die zu perfekter Multikollinearität führen 165

Abb. 20: Geometrische Darstellung einer Regression mit Multikollinearität 169

Verzeichnis der ständig wiederkehrenden Symbole

A_i	Anzahl der Rechnungen mit einer Verweilzeit von i Perioden in der Grundgesamtheit
a_i	Anzahl der Rechnungen mit einer Verweilzeit von i Perioden in der Stichprobe
a_t	Anzahl der in der Periode t bezahlten Rechnungen
B	gesamter Rechnungsbetrag der Grundgesamtheit
b	gesamter Rechnungsbetrag der Stichprobe
B_i	Betrag der Grundgesamtheit mit einer Verweilzeit von i Perioden
b_i	Betrag der Stichprobe mit einer Verweilzeit von i Perioden
$\bar{B}_i$	Durchschnittsbetrag mit einer Verweilzeit von i Perioden pro Rechnung in der Grundgesamtheit
$\bar{b}_i$	Durchschnittsbetrag mit einer Verweilzeit von i Perioden pro Rechnung in der Stichprobe
B_i	Zufallsvariable "Nach i Perioden eingehender Betrag einer Rechnung"
b_{ij}	Ausprägung von B_i bei der Rechnung j
c_i	in Anspruch genommener Skontosatz bei Zahlung nach i Perioden
E_t	Zufallsvariable "Einnahmen in Periode t aus Umsätzen"
e_t	Ausprägung von E_t
$\mathbf{e}$	Vektor der Periodeneinnahmen e_t
F_t	Zufallsvariable "Prognosefehler der Periode t"
f_t	Ausprägung von F_t
i	Verweilzeit in Perioden
$\mathbf{k}$	kumuliertes Spektrum
k_i	Komponente von $\mathbf{k}$ mit der Verweilzeit i

M	Anzahl der Rechnungen in der Grundgesamtheit
m	Anzahl der Rechnungen in der Stichprobe
m_t	Anzahl der in Periode t erstellten Rechnungen
$m_{i,min}$	erforderlicher Stichprobenumfang zur Ermittlung der Komponente i der Verweilzeitverteilung
m_{min}	erforderlicher Stichprobenumfang zur Ermittlung der Verweilzeitverteilung
n	größte Verweilzeit
$\mathbf{p}$	Anzahlverteilung
P_i	Komponente der Anzahlverteilung der Grundgesamtheit mit der Verweilzeit i
p_i	Komponente der Anzahlverteilung der Stichprobe mit der Verweilzeit i
R	Zufallsvariable "Rechnungsbetrag"
r_j	Ausprägung von R
$r_{j,t-i}$	Rechnungsbetrag der j-ten in Periode t-i erstellten Rechnung
$\bar{R}$	durchschnittlicher Rechnungsbetrag in der Grundgesamtheit
$\bar{r}$	durchschnittlicher Rechnungsbetrag in der Stichprobe
$\bar{r}_t$	durchschnittlicher Betrag der in t erstellten Rechnungen
$\mathbf{s}$	Spektrum
s_i	Komponente von $\mathbf{s}$ mit der Verweilzeit i
t	Periodenindex
t	Wert der t-Verteilung
$\mathbf{U}$	Umsatzmatrix
$\mathbf{u}$	Vektor der Periodenumsätze
u_t	Umsatz der Periode t

$\mathbf{w}$	Betragsverteilung
W_i	Komponente der Betragsverteilung der Grundgesamtheit mit der Verweilzeit i
w_i	Komponente der Betragsverteilung der Stichprobe mit der Verweilzeit i
X_i	Null-Eins-Variable
x_{ij}	Ausprägung von X_i
z	Wert der Standardnormalverteilung
$\doteq$	bei großer Anzahl von Ereignissen ungefähr gleich
$\stackrel{!}{=}$	soll sein gleich

1 Einführung

Ein Unternehmen, das seinen Zahlungsverpflichtungen nicht nachkommen kann, ist in seiner Existenz bedroht. Die Prognose der Einnahmen und Ausgaben besitzt deshalb einen hohen Stellenwert. Sie soll unliebsame Entwicklungen rechtzeitig erkennen helfen, damit gegebenenfalls ein planvoller Einsatz von Gegenmaßnahmen eingeleitet werden kann und improvisierte Entschlüsse, die bei überraschend eintretender Zahlungsunfähigkeit notwendig sind, möglichst vermieden werden[1].

Die Prognose der rein finanzwirtschaftlich determinierten Zahlungen ist dabei relativ unproblematisch, da diese gewöhnlich nach Höhe und Fälligkeit recht genau im voraus bekannt sind[2]. Schwieriger ist die Prognose derjenigen Zahlungen, deren Verursachung im Güterbereich liegt und die in der Regel den Hauptanteil des Zahlungsvolumens ausmachen[3]. In besonderem Maße gilt das für die Einnahmen. Während die güterwirtschaftlich determinierten Ausgaben weitgehend dem unmittelbaren Einfluß des Unternehmens unterliegen und damit vergleichsweise sichere Prognosen gestatten[4], deren Problem eher in der Bewältigung der Menge der zu verarbeitenden Informationen zu finden ist, wird besonders der Einzahlungszeitpunkt der Einnahmen aus Umsätzen (im folgenden der Einfachheit halber auch als Umsatzeinnahmen bezeichnet) durch Entscheidungen der Kunden festgelegt.

Eine Prognose der Umsatzeinnahmen muß, will sie Erfolg versprechen, das Zahlungsverhalten der Kunden berücksichtigen. Unter gewissen, noch näher zu bezeichnenden Voraussetzungen kann sie

---

1 Vgl. zum Primat der Planung gegenüber der Improvisation: H. Koch: Betriebliche Planung - Grundlagen und Grundfragen der Unternehmungspolitik, Wiesbaden 1961, S. 12.

2 Vgl. O. Giese: Das betriebswirtschaftliche Problem der Liquidität, Diss. Berlin 1971, S. 233.

3 Ebenda, S. 232.

4 Vgl. H. Lange: Die Vorausbestimmung des Absatzes als Grundlage der unternehmerischen Planung - Dargestellt am Beispiel einer Handelsauskunftei, Diss. Erlangen-Nürnberg 1968, S. 71.

sich dazu der Prognose mit Hilfe von Verweilzeitverteilungen bedienen. Dieses Verfahren erfreut sich in den letzten Jahren einer zunehmenden Beachtung in der Literatur[1]. Im Vordergrund steht dabei seine Einordnung in die Liquiditätspolitik. Darüber darf jedoch nicht vergessen werden, daß die Untersuchung der theoretischen Eigenschaften des Verfahrens noch nicht abgeschlossen ist. Das gilt zunächst einmal für die Ermittlung der Parameter der Verweilzeitverteilung. Hierzu existiert zwar ein theoretisch weitgehend ausgebautes Verfahren[2], es ist jedoch sehr umständlich in der praktischen Handhabung, während für ein anderes[3], in der Anwendung einfacheres Verfahren die theoretische Behandlung noch aussteht. Wenig bekannt ist zum anderen auch über die Genauigkeit der Einnahmeprognose, die bisher nur mit Hilfe von Simulationsexperimenten untersucht worden ist[4].

Die vorliegende Arbeit will die bestehenden Lücken schließen helfen. Sie untersucht dazu schwerpunktartig die Modelleigenschaften dieses Prognoseverfahrens und die Möglichkeiten und Grenzen, die sich daraus für die Ermittlung und Überwachung

---

1 Vgl. z. B.: L. Engels: Zur Optimierung einer zeitabhängigen Liquiditätspolitik bei Saisonunternehmungen, Diss. Köln 1970, S. 72 ff; O. Giese: a. a. O., S. 238 ff; D. Kaiser: Optimale Liquiditätsplanung mit Hilfe der elektronischen Datenverarbeitung und ihre Anwendung in Brauereiunternehmen, Diss. München 1973, S. 115 ff; H.A. Schiel: Die Liquiditätssteuerung als kybernetisches System, Diss. München 1972, S. 132 ff; E. Witte u. H. Klein: Finanzplanung der Unternehmung/Prognose und Disposition, Reinbek 1974, S. 92 ff.

2 Vgl. L. Neumeyer: Zum Problem der Ermittlung von Liquidationsspektren und ihrer Stabilität im Zeitablauf, Diss. Mannheim 1966, S. 23 f.

3 Vgl. H. Langen, R. Edin, G. Kockelkorn, H. Schmitt u. F. Weinthaler: Unternehmensplanung mit Verweilzeitverteilungen - Eine Anleitung für Praktiker, Berlin 1971, S. 27 f.

4 Vgl. R. Edin: Wirkungsweise und Voraussetzungen der Prognose mittels Verweilzeitverteilungen - Ergebnisse einiger Experimente mit Zufallszahlen, ZfB 38 (1968), S. 743 ff.

der Parameter der Verweilzeitverteilung ergeben. Dabei wird von denjenigen Annahmen über das Zahlungsverhalten der Kunden ausgegangen, die üblicherweise bei der Prognose mittels Verweilzeitverteilungen verwendet werden.

Im z w e i t e n Kapitel erfolgt zunächst eine Darstellung der für den weiteren Verlauf benötigten Grundlagen.

Im d r i t t e n Kapitel wird die Ermittlung der Verweilzeitverteilung behandelt. Im Vordergrund steht dabei die Frage nach einem möglichst anwendungsfreundlichen Stichprobenverfahren.

Das v i e r t e Kapitel befaßt sich mit der Genauigkeit der Einnahmeprognose mit Hilfe von Verweilzeitverteilungen. Hier wird zunächst untersucht, welche Eigenschaften der Prognosefehler unter den getroffenen Modellannahmen aufweist und wie eine Intervallprognose der Umsatzeinnahmen aussehen könnte. Daran schließt sich die Betrachtung einiger Maße zur nachträglichen Beurteilung der Prognosegüte an.

Die im f ü n f t e n Kapitel behandelte Überwachung der Stabilität der Verweilzeitverteilung baut auf den Überlegungen des vorherigen Abschnitts auf. Es wird untersucht, welche Möglichkeiten der Überwachung sich gemäß den aus den Annahmen abgeleiteten Eigenschaften des Modells ergeben.

Im s e c h s t e n Kapitel wird geprüft, ob sich unter den Modellannahmen der Prognose mittels Verweilzeitverteilungen auch die multiple Regressionsanalyse zur Prognose der Einnahmen aus den Umsätzen heranziehen läßt.

Die Schlußbetrachtung geht schließlich nach einer kurzen Zusammenfassung noch einmal auf die verwendeten Prämissen ein.

2 Einnahmeprognose mit Hilfe von Verweilzeitverteilungen

2.1 Vorbemerkungen

Das Betriebsgeschehen läßt sich nach LANGEN durch ein System von aufeinanderfolgenden Phasen veranschaulichen, welches je nach Geschäftszweig und Zweck der Untersuchung verschieden tief gegliedert werden kann[1]. Bei der Prognose der Liquidität stehen die in Abbildung 1 wiedergegebenen Phasenfolgen im Vordergrund[2].

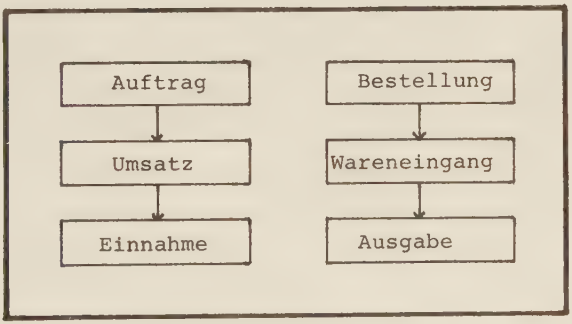

Abb. 1: Phasenfolgen

Die Pfeile symbolisieren die Übergänge zwischen den Phasen. Hier spielen sich die Aktivitäten ab, z. B. die Produktion und die Bezahlung von Lieferantenrechnungen, während die Phasen als Ereignisse aufgefaßt werden, die den Beginn oder den Abschluß der Tätigkeiten angeben[3].

---

1 Vgl. H. Langen: Betriebliche Zahlungsströme und ihre Planung in dynamischer Sicht, ZfB 35 (1965), S. 265 ff; ders.: Preis und konditionspolitische Maßnahmen in der Rezession, in: H. Jacob (Hrsg.): Schriften zur Unternehmensführung, Band 1: Unternehmenspolitik bei schwankender Konjunktur, Wiesbaden 1967, S. 43 ff.

2 Vgl. S. Gahse: Liquiditätsprognosen auf der Grundlage von Phasenfolgen (mit Hilfe von EDVA), Diss. Mannheim 1967, S. 2.

3 Vgl. H. Langen et al., a.a.O., S. 10.

Die durch die Pfeile verdeutlichte Richtung der Abhängigkeit zwischen den Ereignissen wird man in vielen Unternehmen feststellen können, sie ist aber nicht die einzig mögliche. Besonders bei der Auftragsfertigung können erhebliche Wechselwirkungen zwischen der laufenden Produktion und dem Liquiditätsbereich auftreten, denen die Phasenfolgen der Abbildung 1 nicht gerecht werden[1]. Auf diese Fälle soll jedoch nicht eingegangen werden.

Die Einnahmen, die sich aus der Umsatztätigkeit ergeben, können nun auf verschiedene Weise prognostiziert werden[2]:

1. Aufgrund von Erfahrung und Intuition lassen sich s u b j e k t i v e S c h ä t z u n g e n der zukünftigen Einnahmen abgeben. Es gibt Heuristiken, die im Absatzbereich sehr erfolgreich sind[3]. Über ihre Bewährung bei der Einnahmeprognose ist jedoch bisher wenig bekannt[4].

2. Eine weitere Möglichkeit bietet die Z e i t r e i h e n e x t r a p o l a t i o n . Man versucht, eine Gesetzmäßigkeit in der Entwicklung der vergangenen Einnahmen zu erkennen und diese in die Zukunft fortzuschreiben. Hierunter sind die bekannten Verfahren der gleitenden Mittelwerte, des exponentiellen Glättens und der Trendrechnung einzuordnen.

3. Die Schwäche der unter 1. und 2. genannten Verfahren liegt darin, daß sie die Kausalbeziehung zwischen den Umsätzen und den daraus folgenden Einnahmen nur ungenügend bzw. gar

---

1 Vgl. K. R. Thiemann: Das Liquiditätsproblem in der leistungswirtschaftlich orientierten Planung unter besonderer Berücksichtigung der Auftragsfertigung, Diss. Hamburg 1972, S. 48 ff.

2 Vgl. E. Witte u. H. Klein: a.a.O., S. 70 ff; S. Gahse, a.a.O., S. 12 ff; D. Kaiser, a.a.O., S. 127 ff.

3 Vgl. H. J. Rogge: Methoden und Modelle der Prognose aus absatzwirtschaftlicher Sicht - Ein Beitrag zur Prognoseforschung im Unternehmensbereich, Berlin 1972, S. 31.

4 Vgl. E. Witte u. H. Klein, a.a.O., S. 73.

nicht nutzen. Hier setzt die Verweilzeitverteilung an. Sie gibt die Art des Übergangs zwischen den Phasen wieder und ermöglicht die Berücksichtigung der Umsätze bei der Prognose der Einnahmen. Damit gelangt man nicht nur in die glückliche Situation, kausale Prognosen abgeben zu können, die einen hohen theoretischen Stellenwert genießen[1], man schließt auch unmittelbar an die Erfolgsplanung an. Die Einnahmeprognose steht damit nicht mehr frei im Raum, sondern wird Bestandteil der betrieblichen Gesamtplanung[2].

2.2 Begriff der Verweilzeitverteilung

Zur Definition der Verweilzeitverteilung sei von folgendem Beispiel ausgegangen: In der Vergangenheit habe es sich gezeigt, daß die Tagesumsätze stets zu 50% am selben Tag, zu 30% am folgenden Tag, zu 15% am dritten Tag und zu 5% am 4. Tag in bar eingehen[3]. Diese Art der Beschreibung des Zahlungsverhaltens der Kunden ist recht umständlich und läßt sich wesentlich übersichtlicher als Vektor

(50% 30% 15% 5%)

bzw.

(0,50 0,30 0,15 0,05)

wiedergeben.

---

1 Vgl. L. Opitz: Prognosen in der Marktforschung - Die Eignung der demoskopischen Marktforschung zur Formulierung von Prognosen, Wiesbaden 1969, S. 39.

2 Vgl. H. Koch, a.a.O., S. 69.

3 Vgl. zum folgenden H. Ludwig: Die Einnahmeseite des Finanzplans - Eine Budgetstudie, Die Betriebswirtschaft 2 (1932), S. 33 ff; D. Kunkel: Grundlagen der betrieblichen Gesamtplanung (Ein Versuch zu einer Theorie der laufenden Gesamtplanung im Industriebetrieb), Diss. Berlin 1960, S. 74 ff; W. Lücke: Finanzplanung und Finanzkontrolle, Wiesbaden 1962, S. 38 ff; H. Langen: Die Prognose von Zahlungseingängen - Die Abhängigkeit der Bareinnahmen von Umsätzen und Auftragseingängen in dynamischer Betrachtung, ZfB 34 (1964), S. 293; ferner die analoge Betrachtung bei H. Benishay: Managerial Controls of Accounts Receivable: A Deterministic Approach, Journal of Accounting Research 3 (1965), S. 114 ff.

Ein solcher Vektor wird als "Spektralvektor der Liquidation"[1] oder einfach als "Spektralvektor"[2] bezeichnet. Der Spektralvektor gibt an, mit welcher relativen Häufigkeit Verweilzeiten beobachtet werden, wobei die Verweilzeit hier jene Zeit ist, die zwischen Umsatz und Einnahme aus diesem Umsatz verstreicht: Im Beispiel können Forderungen aus Umsätzen 0, 1, 2 oder 3 Tage verweilen, ehe sie beglichen werden. Die relativen Häufigkeiten dieser Verweilzeiten betragen

$$(0{,}50 \quad 0{,}30 \quad 0{,}15 \quad 0{,}05) \ .$$

Der Vektor wird deshalb auch "Verweilzeitverteilung"[3] genannt.

In allgemeiner Schreibweise lautet er

$$\mathbf{s} = (s_0 \ \ s_1 \ \ \dots \ \ s_n) \ .$$

Verteilungen dieser Art lassen sich auch für andere Phasenfolgen ermitteln, etwa für die Folgen "Anfrage - Auftrag" oder "Auftrag - Umsatz" oder auch für die Folge "Auftrag - Einnahme", in der die Phase "Umsatz" nicht explizit berücksichtigt wird. Allgemein kann man definieren:

Die Verweilzeitverteilung gibt an, welcher Teil einer vorangehenden Phase nach einer bestimmten Anzahl von Perioden in eine Folgephase übergeht[4].

Im weiteren Verlauf soll der Übergang zwischen Umsätzen und Einnahmen im Vordergrund stehen, da er für die Liquiditätsprognose von besonderer Bedeutung ist. Im Mittelpunkt der Betrachtung steht also das Liquidationsspektrum. Die dabei getroffenen Aussagen lassen sich jedoch weitgehend auch auf die Verweilzeitverteilung zwischen anderen Phasen übertragen.

---

1 H. Langen: Die Prognose ..., a.a.O., S. 293.

2 Ebenda.

3 H. Langen: Einige Werkzeuge betrieblicher Konjunkturtheorie, ZfB 37 (1967), S. 559.

4 Vgl. H. Langen et al., a.a.O., S. 19.

2.3 Aufbau und Darstellung des Liquidationsspektrums

Wie jede Verweilzeitverteilung hat das Liquidationsspektrum eine Zeit- und eine Mengendimension[1].

Die zeitliche Dimension wird durch die Anzahl der Elemente ausgedrückt, wobei die zeitliche Ordnung durch die Reihenfolge angegeben wird. Da die Liquidation der Umsätze über mehrere Perioden betrachtet wird, ist die Verweilzeitverteilung der dynamischen Analyse zuzurechnen[2].

Die Mengendimension äußert sich im numerischen Wert der einzelnen Elemente. Gemäß der allgemeinen Definition der Verweilzeitverteilung geben diese Werte des Liquidationsspektrums an, welcher Anteil des Periodenumsatzes in einer bestimmten Periode liquidiert wird. Die Summe der Werte ist 1, wenn der gesamte Periodenumsatz in voller Höhe eingeht. Man spricht dann von einem vollständigen oder volldeckenden Spektrum[3]. In der Praxis ist die Summe häufig kleiner als 1. Das kann darauf zurückzuführen sein, daß durch Inanspruchnahme von Skonti durch die Kunden oder durch nachträglich gewährte Preisreduzierung Umsatzforderungen nicht in voller Höhe beglichen werden[4]. Es ist aber auch möglich, daß Teile des Periodenumsatzes, etwa wegen Zahlungsunfähigkeit eines Kunden, überhaupt nicht beglichen werden, wodurch ebenfalls Einzahlungslücken entstehen[5]. Solche Spektren, die eine nur unvollständige Überführung von Umsätzen

---

1 Vgl. H. Kossbiel: Die Umsatzeinnahmen als Gegenstand der unternehmerischen Liquiditätsplanung und Liquiditätspolitik, Berlin 1968, S. 21 f.

2 Vgl. R. Frisch: On the Notation of Equilibrium and Disequilibrium, The Review of Economic Studies 3 (1935/36), S. 100 ff.

3 Vgl. H. Kossbiel, a.a.O., S. 46.

4 Vgl. H. Langen: Die Prognose ..., a.a.O., S. 312 f.

5 Ebenda, S. 317.

in Einnahmen kennzeichnen, werden als unvollständige oder unterdeckende Spektren bezeichnet[1].

Schließlich ist der Fall denkbar, daß die Summe der Werte größer ist als 1. Diese als überdeckend[2] bezeichneten Spektren entstehen hauptsächlich, wenn Kunden wegen Überschreitung von Zahlungsfristen Verzugszinsen oder Konventionalstrafen entrichten müssen[3].

Da die Summe der Komponenten einer Verweilzeitverteilung angibt, welcher Anteil der Vorphase überhaupt in die Folgephase übergeht, läßt sich die Verweilzeitverteilung auch als Multiplikator auffassen[4]. Ein Spektralvektor der Art (0,4 0,3 0,15 0,10) sagt z. B. aus, daß die Umsätze zu 95% liquidiert werden. Er ist ein Umsatzmultiplikator mit dem Wert 0,95.

Betrachtet man die innere Struktur eines Liquidationsspektrums

$$\mathbf{s} = (s_0 \quad s_1 \quad \ldots \quad s_n) ,$$

so lassen sich bei gegebener Periodenlänge, Anzahl der Elemente und Summe der Werte dieser Elemente die drei idealtypischen Fälle

- ausgeglichenes Spektrum: $s_0 = s_1 = \ldots = s_n$,
- monoton fallendes Spektrum: $s_0 \geq s_1 \geq \ldots \geq s_n$ und
- monoton steigendes Spektrum: $s_0 \leq s_1 \leq \ldots \leq s_n$

unterscheiden[5]. Bei einem Vergleich der drei idealtypischen Strukturen schneidet unter sonst gleichen Bedingungen das monoton fallende Spektrum am günstigsten ab, da die Einnahmen

---

1 Vgl. H. Kossbiel, a.a.O., S. 47.

2 Ebenda, S. 48.

3 Ebenda.

4 Vgl. H. Langen: Grundzüge einer betriebswirtschaftlichen Dispositionsrechnung, ZfB 36 (1966), Ergänzungsheft 1, S. 74; S. Pönninghaus: Betriebswirtschaftliche Multiplikatoren, ZfbF 19 (1967), S. 659 ff.

5 Vgl. H. Kossbiel, a.a.O., S. 33 ff.

hier am schnellsten hereinkommen[1]. Empirisch ermittelte Verweilzeitverteilungen werden sich allerdings häufig keinem der drei Typen zuordnen lassen. Will man auch in diesem Fall eine Aussage darüber machen, welches von zwei vorliegenden Spektren die günstigere Struktur im Hinblick auf die Einnahmeentwicklung aufweist, so ist es zweckmäßig, die sukzessiv kumulierten Spektren zu vergleichen[2]. Man stellt dabei für jede Verweilzeit fest, welche der beiden Verteilungen einen höheren kumulierten Wert aufweist, und schreibt dieser Verteilung die Differenz zwischen den kumulierten Werten der betrachteten Verweilzeit gut. Dasjenige Spektrum, für welches sich die höhere Summe dieser Differenzen ergibt, ist dann das günstigere. Zum Beispiel führen die Spektren s_I = (0,3 0,1 0,5 0,1) und s_{II} = (0,2 0,3 0,3 0,2) zu den kumulierten Spektren s_{Ik} = (0,3 0,4 0,9 1,0) und s_{IIk} = (0,2 0,5 0,8 1,0). s_{Ik} ist zweimal um 0,1 günstiger als s_{IIk}, während s_{IIk} einmal um 0,1 günstiger ist als s_{Ik}. s_{Ik} ist deshalb höher einzuschätzen.

Besonders bei langen Verweilzeitverteilungen, d. h. solchen mit einer großen Anzahl von Komponenten, ist es häufig schwierig, sich mit einem Blick eine Vorstellung über die Struktur zu verschaffen. Wesentlich einfacher gelingt das mit Hilfe einer graphischen Darstellung. In Abbildung 2 wird die Verweilzeitverteilung (0,04 0,11 0,09 0,18 0,10 0,13 0,07 0,06 0,04 0,05 0,03 0,04 0,03) durch ein Histogramm anschaulich wiedergegeben.

---

1 Vgl. H. Kossbiel, a.a.O., S. 35.

2 Ebenda, S. 36.

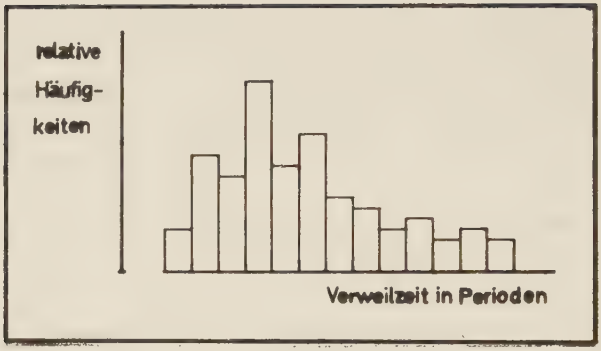

Abb. 2: Darstellung einer Verweilzeitverteilung als Histogramm

Die Werte der Komponenten des Vektors werden hier durch die Flächen der Säulen beschrieben. Wenn man, wie es üblicherweise geschieht, mit Perioden gleicher Länge arbeitet, sind alle Säulen des Histogramms gleich breit. In diesem Fall sind nur noch die Höhen der Säulen von Interesse, und die Verweilzeitverteilung läßt sich als Häufigkeitspolygon darstellen (vgl. Abbildung 3).

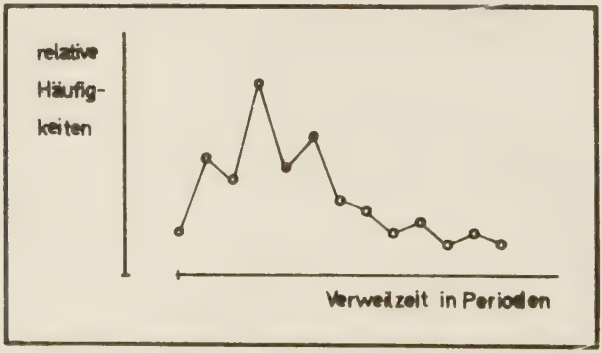

Abb. 3: Darstellung einer Verweilzeitverteilung als Häufigkeitspolygon

Es muß darauf hingewiesen werden, daß die Verbindungslinien lediglich die Funktion der "Augenführung"[1] zwischen den Punkten haben. Die in der Literatur vielfach beobachtbare Praxis, die Punkte durch einen stetigen Kurvenzug zu verbinden[2], ist nicht sehr zweckmäßig. Besonders wenn diese Punkte nicht mehr hervorgehoben werden, läßt sie leicht vergessen, daß es sich hier um die Darstellung eines diskreten Zusammenhangs handelt[3].

2.4 Arten von Verweilzeitverteilungen

Es gibt zwei grundsätzlich verschiedene Arten von Verweilzeitverteilungen: die Mengen- und die Betragsverteilung[4].

Eine Verweilzeitverteilung wird dann als Mengenverteilung bezeichnet, wenn sie angibt, wieviel Prozent des mengenmäßigen Umfangs der vorangegangenen Phase nach einer bestimmten Zeit in die folgende Phase übergeht. Solche Mengenverteilungen lassen sich vorzugsweise im Realgüterbereich[5] verwenden, etwa um den Übergang zwischen zwei Produktionsstufen zu beschreiben. Die Anzahlverteilung, eine Sonderform der Mengenverteilung[6], findet aber auch bei der Betrachtung der Phasenfolge "Umsatz - Einnahme" Verwendung, nämlich wenn gefragt ist, wie viele Rechnungen einer Periode in einer späteren Periode bezahlt werden. Hier handelt es sich dann nicht um einen allgemeinen Mengenbegriff, sondern um die Anzahl abzählbarer Ereignisse.

---

1 Vgl. H. Kossbiel, a.a.O., S. 30.

2 Vgl. z.B. R. Edin u. H. J. Schmitt: Verweilzeitverteilungen und Prognosen - Einige empirische Ergebnisse, ZfbF 21 (1969), S. 486; H. J. Schmitt: Die Planrevision in dynamischen Abgangsfolgemodellen, BFuP 20 (1968), S. 30; H. Langen et al., a.a.O., S. 26.

3 Vgl. H. Kossbiel, a.a.O., S. 30.

4 Vgl. H. Langen et al., a.a.O., S. 22; H. Langen u. F. Weinthaler: Prognose mit Hilfe von Verweilzeitverteilungen, in: P. Mertens (Hrsg.): Prognoserechnung, 2., neubearbeitete u. erweiterte Auflage, Würzburg/Wien 1975, S. 104 ff.

5 Vgl. G. Kockelkorn: Verweilzeitverteilungen und Prognosen im betrieblichen Produktionsbereich, ZfbF 13 (1971), S. 84 ff.

6 Vgl. H. Langen u. F. Weinthaler, a.a.O., S. 105.

Den Mengenverteilungen stehen die Wert- oder Betragsverteilungen gegenüber. Sie geben an, welcher Anteil des Betrages einer Vorperiode nach einer bestimmten Verweilzeit in die nächste Phase übergeht. Der Betrag wird dabei in Geldeinheiten angegeben. Bei einigen Phasenfolgen können durch die Notwendigkeit der Bewertung bei dieser Verteilung erhebliche Schwierigkeiten auftreten. So ist z. B. die Ermittlung einer Betragsverteilung zwischen den Phasen "Kundenanfrage" und "Kundenauftrag" sehr problematisch, wenn der Wert der Anfrage sich erst noch aus Preisverhandlungen ergeben muß oder gar - aufgrund der Möglichkeit mengenmäßiger Unter- oder Überbelieferung - erst nach der Fertigstellung des Auftrags genau feststeht. Der Wert derjenigen Anfragen, die zu keinem Auftrag geführt haben, kann unter solchen Bedingungen nur sehr grob abgeschätzt werden.

Beim Übergang von den Umsätzen zu den Geldeingängen tritt dieses Bewertungsproblem nicht auf. Hier ist der Betrag als Summe der Rechnungsbeträge in der Periode der Rechnungsstellung bekannt, und der Ermittlung einer Betragsverteilung steht unter diesem Gesichtspunkt nichts im Wege.

2.5 Prognose mit dem Liquidationsspektrum

2.5.1 Prognose mit der Betragsverteilung

Das Liquidationsspektrum gibt das in der Vergangenheit beobachtete Zahlungsverhalten wieder[1] und eignet sich zunächst einmal als Diagnoseinstrument[2]. Seine Länge und Struktur informieren darüber, in welchem Umfang die Kunden das maximale Zahlungsziel eingehalten bzw. überschritten haben und wie sich die Zahlungseingänge auf die einzelnen Perioden verteilen.

---

1 Die Ermittlung dieses Spektrums soll an späterer Stelle beschrieben werden. Zunächst sei davon ausgegangen, daß es irgendwie "richtig" als Anzahl- und als Betragsverteilung ermittelt worden ist.

2 Vgl. R. Edin: Dynamische Analyse betrieblicher Systeme - Ein Beitrag zu industriellen Planung, Berlin 1971, S. 180.

Erfolgt diese Betrachtung mit Hilfe der Betragsverteilung, so ist auch die Inanspruchnahme von Skonto seitens der Kunden durch Unterdeckung des Spektrums berücksichtigt. Sie ist in der Regel aber von den übrigen Einnahmeausfällen aus Umsätzen nicht zu trennen. Die Anzahlverteilung kann den Abzug von Skonto nicht aufzeigen. Sie ist in diesem Zusammenhang dennoch von Bedeutung, da sie einen unmittelbaren Einblick in das Verhalten der Kunden gibt, indem sie nur den reinen Zahlungsvorgang erfaßt, während bei der Betragsverteilung eine Überlagerung der Handlung durch die Rechnungsbeträge stattfindet.

Es soll nun an einem Beispiel gezeigt werden, wie mit der Verweilzeitverteilung aus einer Zeitreihe von Umsätzen eine Zeitreihe von Umsatzeinnahmen prognostiziert werden kann.

In der Vergangenheit sei die Betragsverteilung (s_0 s_1 s_2) zwischen Umsätzen und Einnahmen beobachtet worden. Die Einnahmen aus einem Periodenumsatz flossen also jeweils in der Umsatzperiode und den beiden folgenden Perioden zu. Es sei unterstellt, daß diese Verweilzeitverteilung auch in der Zukunft gültig ist, also das Zahlungsverhalten der Kunden für jeden zukünftigen Periodenumsatz widerspiegelt. Man erhält dann die durch Abbildung 4 verdeutlichte Beziehung zwischen Umsätzen und Umsatzeinnahmen.

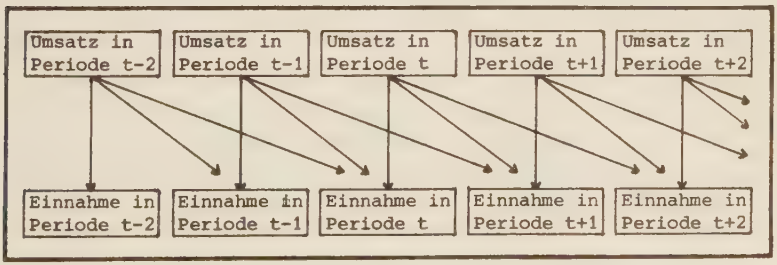

Abb. 4: Zusammenhang zwischen Umsätzen und Umsatzeinnahmen beim Liquidationsspektrum (s_0 s_1 s_2) [1]

1 Vgl. H. Langen: Betriebliche Zahlungsströme..., a.a.O., S. 264.

Der Umsatz der Periode t sei mit u_t bezeichnet und die in t eingehenden Umsatzeinnahmen mit e_t. Unter Verwendung dieser Symbole ergibt sich gemäß Abbildung 4 das folgende Schema zur Berechnung der Einnahmen:

Periode	t-2	t-1	t	t+1	t+2
Umsatz	u_{t-2}	u_{t-1}	u_t	u_{t+1}	u_{t+2}
Einnahme-teile	$u_{t-2}s_0$	$u_{t-1}s_0$	u_ts_0	$u_{t+1}s_0$	$u_{t+2}s_0$
		$u_{t-2}s_1$	$u_{t-1}s_1$	u_ts_1	$u_{t+1}s_1$
			$u_{t-2}s_2$	$u_{t-1}s_2$	$u_t s_2$
Einnahme	.	.	e_t	e_{t+1}	e_{t+2}

Tabelle 1: Rechenschema zur Berechnung der Einnahmen[1]

Die Verbindungslinien in der Tabelle zeigen jeweils an, wie die einzelnen Periodenumsätze liquidiert werden. Aus u_t z. B. resultieren in t Einnahmen in Höhe von u_ts_0, in t+1 solche von u_ts_1 und in t+2 solche von u_ts_2. Die Gesamteinnahmen einer Periode erhält man durch vertikale Addition der Einnahmeteile. Für die Periode t betragen sie

$$e_t = u_ts_0 + u_{t-1}s_1 + u_{t-2}s_2$$

oder

$$e_t = \sum_{i=0}^{2} u_{t-i}s_i \ .$$

1 Vgl. S. Gahse: Eine Systematik der Vorhersageverfahren, IBM Form 81 564, S. 5.

Löst man sich vom Beispiel, indem man das Liquidationsspektrum $(s_0\ s_1\ \dots\ s_n)$ verwendet, so erhält man die allgemeine Gleichung für die Berechnung der Umsatzeinnahmen:

$$e_t = u_t s_0 + u_{t-1} s_1 + \dots + u_{t-n} s_n \tag{1}$$

oder

$$e_t = \sum_{i=0}^{n} u_{t-i} s_i \ . \tag{2}$$

Wegen

$$\sum_{i=0}^{n} u_{t-i} s_i = (u_t\ u_{t-1}\ \dots\ u_{t-n}) \begin{pmatrix} s_0 \\ s_1 \\ \vdots \\ s_n \end{pmatrix} \tag{3}$$

läßt sich e_t auch als Skalarprodukt des Zeilenvektors

$$\mathbf{u} = (u_t\ u_{t-1}\ \dots\ u_{t-n}) \tag{4}$$

und des Spaltenvektors

$$\mathbf{s} = \begin{pmatrix} s_0 \\ s_1 \\ \vdots \\ s_n \end{pmatrix} \tag{5}$$

darstellen[1]:

$$e_t = \mathbf{us} \ . \tag{6}$$

Mit Hilfe der Gleichung (2) bzw. (6) kann man nun eine periodengerechte Prognose der Einnahmen vornehmen. Die Periode t sei die Gegenwart. Soll die erste Einnahmeprognose für t erfolgen[2],

---

1 Vgl. H. Langen: Die Prognose..., a.a.O., S. 295.

2 Streng genommen ist die Prognose von e_t überflüssig, da die Perioden als ohne zeitliche Ausdehnung angenommen werden und mit u_t auch e_t vorliegt. Die Einbeziehung der Prognose von e_t ist jedoch aus formalen Gründen zweckmäßig.

so müssen die vergangenen Umsätze bis zur Periode t-n vorliegen, da das Verfahren einen Vorlauf von n Perioden benötigt, wie man aus Tabelle 1 ablesen kann. Wie weit der Prognosehorizont entfernt ist, hängt davon ab, bis zu welcher Periode in der Zukunft die Umsätze vorhergesagt werden können. Falls ein Absatzplan vorhanden ist, wird man zweckmäßigerweise auf ihn zurückgreifen.

Liegen Umsatzschätzungen bis einschließlich u_{t+m} vor, so berechnet sich der Prognosewert für die Einnahmen der Periode t+j (wobei $0 \leq j \leq m$) als

$$\hat{e}_{t+j} = \sum_{i=0}^{n} u_{t+j-i} s_i, \qquad j = 0,1,\ldots,m. \qquad (7)$$

In welchem Ausmaß bei der Einnahmeprognose auf geschätzte Umsätze anstelle bereits realisierter Umsätze zurückgegriffen werden muß, wird besonders in der Matrizenschreibweise deutlich.

Aus dem Umsätzen u_{t-n} bis u_{t+m} läßt sich die folgende Matrix bilden:

$$\begin{pmatrix} u_t & u_{t-1} & u_{t-2} & \cdots & u_{t-n} \\ u_{t+1} & u_t & u_{t-1} & \cdots & u_{t-n+1} \\ u_{t+2} & u_{t+1} & u_t & \cdots & u_{t-n+2} \\ \vdots & \vdots & \vdots & & \vdots \\ u_{t+m} & u_{t+m-1} & u_{t+m-2} & \cdots & u_{t-n+m} \end{pmatrix},$$

die LANGEN als "dynamische Umsatzmatrix"[1] bezeichnet. Die Zeitreihe der prognostizierten Einnahmen ergibt sich durch Multiplikation dieser Matrix mit dem Spaltenvektor **s** als

---

1 Vgl. H. Langen: Die Prognose..., a.a.O., S. 295.

$$\begin{pmatrix} \hat{e}_t \\ \hat{e}_{t+1} \\ \hat{e}_{t+2} \\ \vdots \\ \hat{e}_{t+m} \end{pmatrix} = \begin{pmatrix} u_t & u_{t-1} & u_{t-2} & \cdots & u_{t-n} \\ u_{t+1} & u_t & u_{t-1} & \cdots & u_{t-n+1} \\ u_{t+2} & u_{t+1} & u_t & \cdots & u_{t-n+2} \\ \vdots & \vdots & \vdots & & \vdots \\ u_{t+m} & u_{t+m-1} & u_{t+m-2} & \cdots & u_{t-n+m} \end{pmatrix} \begin{pmatrix} s_0 \\ s_1 \\ s_2 \\ \vdots \\ s_n \end{pmatrix} \quad (8)$$

oder

$$\hat{\mathbf{e}} = \mathbf{U}\mathbf{s} \quad . \qquad (9)$$

In der dem Einnahmeprognosewert zugehörigen Zeile der Umsatzmatrix sind die geschätzten Umsätze jeweils diejenigen, deren Index größer ist als t.

2.5.2 Prognose mit der Anzahlverteilung

Die Prognose der Umsatzeinnahmen mit Hilfe der Anzahlverteilung erfolgt in zwei Schritten. Zuerst wird eine Prognose über die in der Periode erwartete Anzahl von bezahlten Rechnungen erstellt, dann wird von dieser Anzahl auf den Betrag geschlossen. Die Betragsprognose erfolgt also indirekt.

Der Umsatz einer Periode t sei wieder mit u_t bezeichnet. Die Anzahl der in dieser Periode erstellten Rechnungen sei m_t. Der mittlere Betrag der in t ausgehenden Fakturen ergibt sich als

$$\bar{r}_t = \frac{u_t}{m_t} \quad . \qquad (1)$$

Die Anzahl der in t bezahlten Rechnungen, die aus t oder einer früheren Periode stammen, sei a_t.

Die Übergangsfunktion (s_0 s_1 ... s_n) zwischen den Phasen "Umsatz" und "Einnahme" sei jetzt eine Anzahlverteilung, gebe also an, welcher Anteil der Anzahl der in einer Periode erstell-

ten Rechnungen stets eine Verweilzeit von 0,1, ..., n Perioden aufweist. Dabei sei zunächst von einer volldeckenden Verweilzeitverteilung ausgegangen.

Die Prognose erfolgt in Analogie zur direkten Betragsprognose. Die vorhergesagte Anzahl der in t eingehenden Rechnungen setzt sich auch hier aus n + 1 Teilen zusammen:

$$\hat{a}_t = m_t s_0 + m_{t-1} s_1 + \ldots + m_{t-n} s_n \qquad (2)$$

oder

$$\hat{a}_t = \sum_{i=0}^{n} m_{t-i} s_i \,. \qquad (3)$$

$m_{t-i} s_i$ gibt an, wieviel der m_{t-i} in t-i erstellten Rechnungen in t bezahlt werden. Berücksichtigt man, daß in t-i der durchschnittliche Rechnungsbetrag $\bar{r}_{t-i}$ beträgt, so liegt es nahe, in t den aus t-i stammenden Einzahlungsbetrag als $\bar{r}_{t-i} m_{t-i} s_i$ zu schätzen. Die Prognose des gesamten in t eingehenden Betrages ergibt sich deshalb als

$$\hat{e}_t = \sum_{i=0}^{n} \bar{r}_{t-i} m_{t-i} s_i \,. \qquad (4)$$

Wegen (1) ist

$$\bar{r}_{t-i}\, m_{t-i} = u_{t-i}$$

und damit

$$\hat{e}_t = \sum_{i=0}^{n} u_{t-i} s_i \,. \qquad (5)$$

Diese Gleichung ist formal identisch mit der Gleichung der unmittelbaren Betragsprognose des vorigen Abschnitts. Auch die Zeitreihe der prognostizierten Beträge der Einnahmen lautet in der Matrizenschreibweise genau wie dort

$$\hat{\mathbf{e}} = \mathbf{U}\mathbf{s} \,. \qquad (6)$$

Man beachte jedoch, daß die s_i hier Elemente der Anzahlverteilung sind. Sie unterscheiden sich in der Regel von den Elementen der entsprechenden Betragsverteilung. Dadurch ergeben sich für eine Periode verschiedene prognostizierte Beträge, wenn man die Prognose einmal direkt und einmal indirekt durchführt. Geht man von der Annahme einer volldeckenden Verweilzeitverteilung ab und unterstellt man, daß Einnahmeminderungen etwa durch Abzug von Skonto möglich sind, so zeigt sich eine Schwäche des indirekten Ansatzes gegenüber dem direkten: Die Anzahlverteilung kann nur berücksichtigen, ob eine Rechnung überhaupt bezahlt wird oder nicht, Schmälerungen des Betrages schlagen sich jedoch nicht in ihr nieder. Treten solche Einnahmeminderungen tatsächlich auf, so führt die indirekte Betragsprognose zu systematischen Überschätzungen der Einnahmen[1]. Aus dieser Sicht ist die direkte Betragsprognose günstiger zu beurteilen.

Die systematische Überschätzung läßt sich durch explizite Berücksichtigung der Einnahmeminderung, die durch Abzug von Skonto entsteht, ausschalten. Der von den Kunden bei Bezahlung nach i Perioden durchschnittlich in Anspruch genommene Skontosatz sei c_i. Die Einnahmeprognose lautet dann[2]

$$\hat{e}_t = \sum_{i=0}^{n} (1-c_i)\, u_{t-i} s_i \,. \qquad (7)$$

Sind die tatsächlichen Werte von c_i nicht bekannt, so kann man an ihrer Stelle auch die Sätze verwenden, die den Kunden gemäß den Zahlungsbedingungen eingeräumt werden. Diese brauchen allerdings nicht unbedingt mit dem tatsächlichen Zahlungsverhalten der Kunden übereinzustimmen.

---

1 Vgl. R. Edin u. H. J. Schmitt, a.a.O., S. 497 ff. Die beiden Autoren demonstrieren an einem empirischen Beispiel die systematische Überschätzung bei der Prognose mit der Anzahlverteilung. Sie geben allerdings nicht an, nach welcher Formel die Prognosewerte berechnet worden sind, und es kann deshalb nur eine Vermutung bleiben, daß der Grund in der Nichtberücksichtigung von Skonti zu suchen ist.

2 Vgl. R. Edin: Übergangsfunktionen in betriebswirtschaftlichen Systemen, ZfB 39 (1969), S. 576 f.

Auch bei Verwendung der richtigen Werte von c_i liefert die Anzahlverteilung verzerrte Prognosen für die Einnahmen, wenn die Verweilzeit der Rechnungen von der Höhe der Rechnungsbeträge abhängig ist[1]. Trotz dieser Einschränkung hat die Prognose mit der Anzahlverteilung eine gewisse Bedeutung dadurch, daß diese Verteilung erheblich einfacher und genauer ermittelt werden kann als die Betragsverteilung.

2.6 Deterministische und stochastische Prognose mit der Verweilzeitverteilung

Bei der Herleitung der Prognosegleichung

$$\hat{\mathbf{e}} = \mathbf{U}\mathbf{s} \qquad (1)$$

wurde unterstellt, daß sämtliche Periodenumsätze jeweils gemäß der verwendeten Verweilzeitverteilung liquidiert werden. Diese Forderung schließt zwei Überlegungen ein:

1. Das Prognoseverfahren ist deterministisch.
2. Die Verweilzeitverteilung ist eine stationäre Systemfunktion.

Ein deterministisches Prognoseverfahren ist dadurch gekennzeichnet, daß durch den Input der Output vollständig bestimmt wird[2]. Die Verknüpfung von Input und Output enthält kein Zufallselement. Für die Prognosegleichung (1) bedeutet das, daß $\mathbf{s}$ als Vektor mit deterministischen Elementen angesehen wird: Für jede Prognoseperiode wird nur eine ganz bestimmte Liquidationsweise der Umsatzforderungen unterstellt, die durch $\mathbf{s}$ angegeben werden kann.

Stationarität bedeutet im deterministischen System die zeitliche Konstanz des Systemverhaltens[3]. In der Prognosegleichung

---

1 Vgl. H. Langen et al., a.a.O., S. 54.

2 Vgl. R. Edin: Dynamische Analyse..., a.a.O., S. 57.

3 Ebenda, S. 59.

(1) kommt die Annahme der Stationarität dadurch zum Ausdruck, daß stets mit dem gleichen Spektrum **s** gerechnet wird. Unterstellte man Nichtstationarität, so müßte für jede Periode eine eigene Verweilzeitverteilung benutzt werden. Statt des Spaltenvektors **s** wäre dann eine Matrix der Form

$$\mathbf{S} = \begin{pmatrix} s_{0,t} & s_{0,t+1} & s_{0,t+2} & \cdots & s_{0,t+m} \\ s_{1,t-1} & s_{1,t} & s_{1,t+1} & \cdots & s_{1,t+m-1} \\ s_{2,t-2} & s_{2,t-1} & s_{2,t} & \cdots & s_{2,t+m-2} \\ \vdots & \vdots & \vdots & & \vdots \\ s_{n,t-n} & s_{n,t-n+1} & s_{n,t-n+2} & \cdots & s_{n,t+m-n} \end{pmatrix} \quad (2)$$

zu benutzen. Darin gibt z. B. $s_{0,t}$ den Anteil des in t getätigten Umsatzes an, welcher in derselben Periode bezahlt wird.

Die Zeitreihe der prognostizierten Einnahmen findet man dann auf der Hauptdiagonalen der Matrix

$$\hat{\mathbf{E}} = \mathbf{US} \; . \quad (3)$$

Die Matrix **S** zeigt, warum man gewöhnlich auf den stationären Ansatz angewiesen ist. Diejenigen Elemente dieser Matrix, deren Periodenindizes gleich oder kleiner als t sind, lassen sich aus Vergangenheitsdaten ermitteln. Außer diesen Elementen enthält die Matrix aber auch solche mit einem größeren Periodenindex, welche selbst prognostiziert werden müssen. Die Prognose von Verweilzeitverteilungen jedoch gestaltet sich sehr schwierig[1].

Wenn man wegen dieses speziellen Prognoseproblems auf die stationäre Prognosegleichung (1) zurückgreift, so sollte man nicht vergessen, daß die Realitätsnähe des Prognoseverfahrens

---

1 Vgl. zu diesem Problem H. Kossbiel, a.a.O., S. 194 ff; L. Neumeyer, a.a.O., S. 111 ff.

darunter leidet. Eine Reihe von Gründen spricht nämlich für eine Veränderung des Liquidationsspektrums im Zeitablauf. Hier sind zunächst jene Ereignisse zu nennen, welche die äußeren Rahmenbedingungen des wirtschaftlichen Handelns z. T. erheblich beeinflussen können, wie Veränderung der Konjunkturlage oder Änderung der Steuergesetzgebung, aber auch Kriege und Katastrophen[1]. Betreiben Kunden ein saisonabhängiges Geschäft, so kann ihr Zahlungsverhalten aufgrund saisonaler Liquiditätsengpässe ebenfalls zeitabhängig sein[2]. Veränderungen sind ferner möglich, wenn sich die Zusammensetzung des Kundenkreises[3] ändert oder aktive Einwirkung auf das Zahlungsverhalten durch Konditionspolitik[4] oder Mahnwesen betrieben wird.

Wie schnell sich das Liquidationsspektrum aufgrund solcher Einflüsse verändert, kann nicht allgemein angegeben werden. Man wird jedoch davon ausgehen können, daß sich das Zahlungsverhalten der Kunden nicht von heute auf morgen grundlegend ändert, sondern sich eher evolutorisch entwickelt, solange man es nicht massiv durch eigene Maßnahmen beeinflußt[5]. Da die Liquiditätsprognose definitionsgemäß nur einen kurzen Zeitraum umfaßt[6], darf man deshalb hoffen, daß der durch die Anwendung eines stationären anstelle des richtigen nichtstationären Ansatzes verursachte Fehler vertretbar klein bleibt.

---

1 Vgl. H. Langen: Einige Werkzeuge..., a.a.O., S. 560; O. Giese, a.a.O., S. 251.

2 Vgl. D. Koebel: Finanzplanung in der Schuhindustrie, ZfhF, NF, 7 (1955), S. 291 ff; W. Neitzel: Finanzplanung in der Textilindustrie, ZfhF, NF, 7 (1955), S. 294 ff.

3 Vgl. H. Kossbiel, a.a.O., S. 86 ff.

4 Vgl. L. Engels, a.a.O., S. 101.

5 Vgl. H. Langen: Die Prognose..., a.a.O., S. 319; R. Edin: Wirkungsweise..., a.a.O., S. 756.

6 Vgl. L. Orth: Die kurzfristige Finanzplanung industrieller Unternehmungen, Köln/Opladen 1961, S. 57 ff; J. Zaunert: Die Liquidität der Betriebswirtschaft als Problem der Regelung und Steuerung des Zahlungs- und Kreditverkehrs, Diss. Braunschweig 1975, S. 5.

Ist es nun überhaupt zweckmäßig, die Prognosegleichung als einen deterministischen Ansatz zu interpretieren? Ein solcher Ansatz unterstellt eine eindeutige Beziehung zwischen Umsatz und Einnahmen. Wie schlecht es jedoch um die Eindeutigkeit bei der Prognose von Umsatzeinnahmen bestellt ist, zeigt sich bereits, wenn man eine einzelne Rechnung betrachtet. Für eine deterministische Prognose der Bezahlung dieser Rechnung müßte eine Aussage der Art "Dieser Kunde zahlt immer nach genau i Perioden" oder "Dieser Kunde hat sich verpflichtet, nach i Perioden zu zahlen" gemacht werden können.

In der Realität sind solche Fälle eher die Ausnahme. Kommen für den Zahlungseingang einer Rechnung mehrere Perioden in Frage, d. h. sind mehrere Verweilzeiten denkbar, so ist ein deterministischer Ansatz nicht mehr angemessen, da er nur einen Zahlungstermin berücksichtigen kann. Dazu kommt, daß sich die Prognose mit Hilfe von Verweilzeitverteilungen in der oben dargestellten Form nicht mit einzelnen Rechnungsbeträgen und den daraus resultierenden Einnahmen, sondern mit Aggregaten solcher Größen befaßt. Wegen dieser Aggregation ist die Liquidation einzelner Rechnungen nicht mehr zu verfolgen. Greift man aus den für eine bestimmte Periode mit der Verweilzeitverteilung prognostizierten Einnahmen eine DM heraus, so läßt sich nicht angeben, aus welcher Umsatzperiode sie stammt[1]. Anders herum kann man auch nicht sagen, in welcher Folgeperiode eine ganz bestimmte DM einer Umsatzperiode liquidiert wird.

Für die Liquidation einer einzelnen DM kann man jedoch eine Wahrscheinlichkeitsaussage machen. Die relativen Häufigkeiten der Betragsverteilung lassen sich nämlich als Wahrscheinlichkeiten dafür interpretieren, daß eine DM eines Periodenumsatzes eine ganz bestimmte Verweilzeit hat[2]. Nach dieser Auf-

---

1 Vgl. S. Pönninghaus, a.a.O., S. 661.

2 Vgl. H. Langen: Preis- und konditionspolitische Maßnahmen..., a.a.O., S. 51.

fassung geben z. B. die Elemente der Betragsverteilung (0,5 0,3 0,2) die Wahrscheinlichkeiten dafür an, daß die Verweilzeit dieser DM 0,1 oder 2 Perioden beträgt. Entsprechendes gilt auch für die Anzahlverteilung, nur werden hier nicht einzelne "DM-Stücke", sondern ganze Rechnungen betrachtet. Damit erhält die Verweilzeitverteilung den Charakter einer Wahrscheinlichkeitsverteilung.

Schwierigkeiten entstehen bei dieser Interpretation formal dadurch, daß in einer Wahrscheinlichkeitsverteilung die Summe der Einzelwahrscheinlichkeiten stets 1 ist, während Verweilzeitverteilungen nicht immer volldeckend sind. Liegt eine unterdeckende Verweilzeitverteilung vor, so ist die Summe vielmehr kleiner als 1. Man kann hier jedoch eine zusätzliche sehr große Verweilzeit einführen, die bei der Prognose nicht mitverwendet wird, und ihr die Differenz als Wahrscheinlichkeit zuordnen. Nicht lösbar ist dagegen das Problem überdekkender Verweilzeitverteilungen. Bei der weiteren Betrachtung muß dieser Fall deshalb ausgeschlossen werden, wenn wahrscheinlichkeitstheoretische Aussagen gemacht werden. Sicherlich ist das ein Mangel, der aber weniger schwer wiegt, wenn man bedenkt, daß solche Verweilzeitverteilungen zwar vorstellbar, jedoch keineswegs die Regel sind.

Die Interpretation der Verweilzeitverteilung als Wahrscheinlichkeitsverteilung soll sich fortan auch in der Symbolik niederschlagen. Um ferner zwischen den beiden Verteilungsarten besser unterscheiden zu können, seien die Elemente der Betragsverteilung mit w_i bezeichnet:

$$\mathbf{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_i \\ \vdots \\ w_n \end{pmatrix} \qquad (4)$$

und die der Anzahlverteilung mit p_i:

$$\mathbf{p} = \begin{pmatrix} p_1 \\ \vdots \\ p_i \\ \vdots \\ p_n \end{pmatrix} \quad . \tag{5}$$

Die Prognosegleichungen für die Einnahmen lauten damit bei direkter Betragsprognose

$$\hat{e}_t = \sum_{i=0}^{n} u_{t-i}\, w_i \tag{6}$$

bzw. $\hat{\mathbf{e}} = \mathbf{U}\mathbf{w}$ (7)

und bei indirekter Betragsprognose

$$\hat{e}_t = \sum_{i=0}^{n} u_{t-i}\, p_i \tag{8}$$

bzw. $\hat{\mathbf{e}} = \mathbf{U}\mathbf{p}$. (9)

Der Zusammenhang zwischen den Einnahmen und den Umsätzen ist jetzt stochastischer Natur, und die Prognosewerte haben den Charakter von Erwartungswerten. Damit eine gute Übereinstimmung der Prognosen mit den tatsächlich realisierten Einnahmen erzielt wird, müssen einige Voraussetzungen erfüllt sein[1]:

1. Die Anzahl der pro Periode erstellten Rechnungen ist groß.

2. Die Rechnungen werden voneinander unabhängig bezahlt. Wenn nur wenige Großabnehmer existieren, ist das oft nicht der Fall, weil diese Kunden die Rechnungen aus Kostengründen häufig paketweise bezahlen. Die Prämisse könnte somit auch lau-

---

1 Vgl. dazu R. Edin: Wirkungsweise..., a.a.O., S. 748 f.

ten: Es existieren viele Kunden von etwa gleicher Bedeutung. Sollte es tatsächlich einige Großkunden geben, so empfiehlt sich eine getrennte Behandlung bei der Einnahmeprognose.

3. Die verwendete Wahrscheinlichkeitsverteilung **w** bzw. **p** ist über die Zeit hinweg stabil, d. h. das Zahlungsverhalten der Kunden ändert sich nicht. Sofern nicht alle Kunden dasselbe Zahlungsverhalten aufweisen, schließt diese Voraussetzung die Annahme eines gleichbleibend strukturierten Kundenstammes ein.

Dazu kommen noch zwei Prämissen mehr technischer Art:

4. Die ermittelte Verweilzeitverteilung gibt die wahre Verweilzeitverteilung ausreichend genau wieder, ist also repräsentativ für die Zahlungsweise der Kunden[1]. Diese Voraussetzung stellt auf die Art der Ermittlung des Liquidationsspektrums ab.

5. Die Umsätze zukünftiger Perioden lassen sich genügend genau prognostizieren. Liegen Zahlen für die zukünftigen Umsätze vor, so ist darauf zu achten, ob sie auf der gleichen Periodenlänge basieren wie die Liquiditätsprognose. Das braucht insbesondere dann nicht zuzutreffen, wenn die Zahlen dem Absatzplan entnommen werden, bei dessen Erstellung Liquiditätsgesichtspunkte oft nicht berücksichtigt werden. Sind die Perioden der Umsatzprognose kleiner, so lassen sie sich entsprechend zusammenfassen. Wegen des besonderen Anliegens der Liquiditätsprognose, die Zahlungsfähigkeit zu jedem Zeitpunkt sicherzustellen, ist jedoch eher damit zu rechnen, daß die Perioden der Liquiditätsprognose kürzer sind. Man kann sich dann damit behelfen, den vorhergesagten Umsatz einer Absatzperiode zu gleichen Teilen auf die in ihr enthaltenen Perioden der Liquiditätsprognose aufzuteilen[2].

---

1 Vgl. R. Edin: Wirkungsweise..., a.a.O., S. 749.

2 Vgl. H. Langen: Die Prognose..., a.a.O., S. 313.

Über die günstigste Periodenlänge bei der Einnahmeprognose ist eine allgemein verbindliche Aussage nicht möglich. Es stehen sich zwei Aspekte gegenüber. Da die Zahlungsfähigkeit zu jedem Zeitpunkt gewährleistet sein muß, besteht ein großes Interesse an möglichst kurzen Basisperioden. Andererseits verlangt die 1. Prämisse tendenziell lange Basisperioden. Man muß hier einen Kompromiß schließen und solchen Finanzlücken, die innerhalb der im Modell als Zeitpunkt behandelten Perioden auftreten können, durch entsprechende Kassenhaltung begegnen[1].

In der Praxis hat man mit dem Prognosemodell DYPOL von IBM, das mit Verweilzeitverteilungen arbeitet[2], bei einer Periodenlänge von einem Monat recht genaue Liquiditätsprognosen erstellen können[3]. In der Literatur werden häufig kürzere Perioden gefordert, etwa der Arbeitstag oder eine bestimmte Anzahl von Arbeitstagen[4].

---

1 Vgl. D. Kaiser, a.a.O., S. 77 ff.

2 Vgl. S. Gahse: Dynamische Liquiditätsplanung DYPOL auf IBM System/360, IBM Form 8 064 200.

3 Vgl. F. H. Lütticken: Praktische Erfahrungen mit dem Finanzplanungsmodell DYPOL, IBM-Nachrichten 19 (1969), S. 771 ff.

4 Vgl. etwa H. Kossbiel, a.a.O., S. 280 f; L. Neumeyer, a.a.O., S. 16 f.

3 Ermittlung der Verweilzeitverteilung

3.1 Vorbemerkungen

Ehe die eigentliche Ermittlung der Verweilzeitverteilung betrachtet wird, erscheinen einige Vorüberlegungen angebracht. Soll die Verweilzeitverteilung der nachträglichen Beschreibung des Zahlungsverhaltens der Kunden dienen, soll also festgestellt werden, wie die Rechnungen eines vergangenen Zeitraums bezahlt worden sind, so bilden diese Rechnungen im Sinne des Untersuchungszwekkes eine Grundgesamtheit. Das Zahlungsverhalten wird durch die Liquidation aller Rechnungen des betrachteten Zeitraums ausgedrückt. Zur Bestimmung der Verweilzeitverteilung kann man eine Vollerhebung durchführen, also sämtliche Rechnungen erfassen, man kann aber auch eine Stichprobe ziehen und von dieser auf die Grundgesamtheit schließen.

Bei der Prognose mit der Verweilzeitverteilung wird gleichbleibendes durchschnittliches Zahlungsverhalten im Zeitablauf unterstellt. Das Liquidationsspektrum soll nicht nur die Zahlungsweise der Vergangenheit, sondern auch die der Zukunft repräsentieren. Es erscheint deshalb angebracht, die Grundgesamtheit um die Rechnungen desjenigen zukünftigen Zeitraums zu erweitern, in dem das Zahlungsverhalten der Kunden dem der Vergangenheit entspricht.

Abbildung 5 soll die Grundgesamtheit bei der Prognose verdeutlichen.

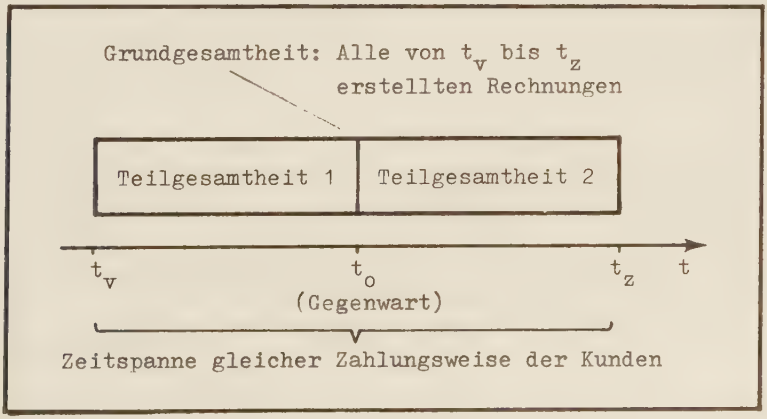

Abb. 5: Grundgesamtheit bei der Ermittlung der Verweilzeitverteilung zu Prognosezwecken

Die Grundgesamtheit ist endlich. Da man im voraus nicht weiß, wie lange das Zahlungsverhalten unverändert bleibt, ist ihr Umfang aber nicht bekannt. Er soll im folgenden in Übereinstimmung mit den Prämissen der Prognose mit Hilfe des Liquidationsspektrums als sehr groß angenommen werden.

Die Ermittlung der Verweilzeitverteilung kann nur aus der Teilgesamtheit 1 erfolgen. Die vollständige Erhebung dieser Teilgesamtheit läßt sich als Entnahme einer Stichprobe maximalen Umfangs interpretieren. Die Ziehung einer Stichprobe erfolgt somit nicht aus der gesamten Grundgesamtheit, sondern aus einer größeren übergeordneten Stichprobe nach Art der zweiphasigen Auswahl[1]. Da die Teilgesamtheit 1 repräsentativ für die Grundgesamtheit ist, darf die Stichprobe jedoch genauso behandelt werden, als stamme sie direkt aus der Grundgesamtheit[2].

3.2 Ermittlung der Anzahlverteilung

Der Ermittlung der Anzahlverteilung wird das folgende Modell zugrundegelegt:
Die Grundgesamtheit umfaßt M Rechnungen, wobei M unbekannt ist, da man noch nicht weiß, wieviele Rechnungen in der Zukunft erstellt werden. Die Anzahl der Rechnungen mit einer Verweilzeit von i (i = 0, ..., n) beträgt A_i. Die Komponente P_i der Anzahlverteilung der Grundgesamtheit ergibt sich somit als

$$P_i = \frac{A_i}{M} , \qquad i = 0, \ldots, n . \qquad (1)$$

---

1 Vgl. Stat. Bundesamt (Hrsg.): Stichproben in der amtlichen Statistik, Stuttgart/Mainz 1960, S. 43 f; G. Menges u. H.J. Skala: Grundriß der Statistik, Teil 2: Daten - Ihre Gewinnung und Verarbeitung, Opladen 1973, S. 118 ff.

2 Vgl. F. Yates: Sampling Methods for Census and Surveys, 3rd edition, 2nd impression, London 1965, S. 171.

Zieht man eine Zufallsstichprobe von m Rechnungen und findet darin a_i Rechnungen mit der Verweilzeit i, so erhält man mit

$$\hat{P}_i = p_i = \frac{a_i}{m} \quad , \qquad i = 0, \ldots, n \qquad (2)$$

einen erwartungstreuen Schätzwert für P_i.

Für eine Intervallschätzung von P_i benötigt man außer P_i noch die Varianz von $\hat{P}_i$. Sie lautet bei Entnahme ohne Zurücklegen und einem Stichprobenumfang von m

$$V(\hat{P}_i) = \frac{P_i\ (1-P_i)}{m} \frac{M-m}{M-1} \quad , \qquad (3)$$

ist unter anderem also auch vom unbekannten Umfang M der Grundgesamtheit abhängig. Im folgenden soll M als sehr groß im Vergleich zu m angenommen werden[1]. Das Entnahmemodell ohne Zurücklegen kann dann durch das mit Zurücklegen approximiert werden und die Varianz von $\hat{P}_i$ beträgt

$$V(\hat{P}_i) = \frac{P_i\ (1-P_i)}{m} \quad . \qquad (4)$$

Sie läßt sich erwartungstreu schätzen als

$$\hat{V}(\hat{P}_i) = \frac{p_i\ (1-p_i)}{m-1} \quad . \qquad (5)$$

Zöge man nun aus der Grundgesamtheit fortgesetzt mit Zurücklegen Stichproben vom Umfang m, so wäre $\hat{P}_i$ annähernd normalverteilt, wenn m genügend groß ist. Die Approximation ist dabei umso besser, je größer m ist und je näher p_i bei 0,5 liegt. Eine Faustregel besagt, daß das kleinere der beiden Produkte mp_i bzw. $m(1-p_i)$ mindestens den Wert 10 aufweisen sollte[2].

---

1 Diese Annahme entspricht der 1. Prämisse der Prognose mit der Verweilzeitverteilung, die eine große Anzahl von Rechnungserstellungen pro Periode verlangt.

2 Vgl. W. L. Hays: Statistics for the Social Sciences, 2nd edition, London 1974, S. 308.

Manche Autoren sehen auch schon einen Wert von 5 als aureichend an[1].

Bei ausreichender Approximation durch die Normalverteilung läßt sich für P_i gemäß

$$p_i - t\sqrt{\hat{V}(\hat{P}_i)} \leq P_i \leq p_i + t\sqrt{\hat{V}(\hat{P}_i)} \qquad (6)$$

die Intervallschätzung

$$p_i - t\sqrt{\frac{p_i(1-p_i)}{m-1}} \leq P_i \leq p_i + t\sqrt{\frac{p_i(1-p_i)}{m-1}} \qquad (7)$$

abgeben. Dabei wird t der entsprechenden t-Verteilung mit m-1 Freiheitsgraden entnommen. Solange m nicht kleiner als 30 ist, macht man aber auch keinen nennenswerten Fehler, wenn man bei einem Signifikanzniveau von $\alpha = 0{,}05$ grundsätzlich mit $t = 2$ rechnet[2].

Das nach (7) berechnete Konfidenzintervall für P_i wird symmetrisch um p_i angeordnet. Bei sehr kleinen P_i, wie sie im hinteren Teil von Liquidationsspektren oft vorkommen[3], sind die Verteilungen der zugehörigen Schätzfunktionen $\hat{P}_i$ stark linkssteil. Arbeitet man hier mit einem symmetrischen Konfidenzintervall, so werden beide Intervallgrenzen zu klein gewählt. Der Stichprobenwert p_i liegt dann nämlich mit einer geringeren Wahrscheinlichkeit als $\alpha/2$ unterhalb der unteren Intervallgrenze $p_i - t\sqrt{\hat{V}(\hat{P}_i)}$ und übersteigt mit einer größeren Wahrscheinlichkeit als $\alpha/2$ die obere Grenze $p_i + t\sqrt{\hat{V}(\hat{P}_i)}$ [4].

---

1 Vgl. Th. H. Wonnacott u. R. J. Wonnacott: Introductory Statistics for Business and Economics, New York 1972, S. 130.

2 Vgl. z. B. W. L. Hays, a.a.O., S. 885.

3 Vgl. etwa H. Langen: Die Prognose..., a.a.O., S. 300; L. Neumeyer, a.a.O., S. 87; R. Edin u. H. J. Schmitt, a.a.O., S. 490.

4 Vgl. W. G. Cochran: Stichprobenverfahren, Berlin/New York 1972, S. 59.

In solchen Fällen empfiehlt es sich, asymmetrische Konfidenzintervalle zu wählen, wie sie in Nomogrammen wiedergegeben werden, die in entsprechenden Tabellenwerken[1] oder dem Anhang vieler Lehrbücher[2] zu finden sind.

3.3 Bestimmung des Mindeststichprobenumfangs zur Ermittlung der Anzahlverteilung

Bei der Intervallschätzung der Komponenten der Anzahlverteilung ergeben sich mit zunehmendem Stichprobenumfang bessere Schätzergebnisse einmal durch die stärkere Annäherung an die Normalverteilung, zum anderen durch die Verringerung der Varianzen $V(\hat{P}_i)$ und die daraus resultierende Verkleinerung der Intervalle. Da mit der Erhöhung der Schätzgenauigkeit der Erhebungsaufwand überproportional ansteigt, wird man nicht die höchstmögliche Genauigkeit anstreben, sondern einen Kompromiß suchen. Man könnte etwa fordern:

Es sollen nur so viele Rechnungen in die Stichprobe gelangen, daß mit einer Wahrscheinlichkeit von $(1 - \alpha)$ die Intervallgrenzen um nicht mehr als den Betrag e nach oben und unten von P_i abweichen. Erscheint der nach dieser Vorschrift errechnete Stichprobenumfang vom Arbeitsaufwand her als zu groß, sollte überprüft werden, ob man sich nicht mit einem niedrigeren Genauigkeitsniveau zufriedengeben und einen entsprechend höheren Wert für e vorgeben will, für den dann abermals der erforderliche Stichprobenumfang bestimmt wird. Auf diese Weise läßt sich nach einigen Schritten ein Stichprobenumfang ermitteln, der sowohl Kosten- als auch Genauigkeitsüberlegungen Rechnung trägt.

---

1 Vgl. z. B. W. H. Beyer (ed.): CRC Handbook of Tables for Probability and Statistics, 2nd edition, Cleveland 1968, S. 220 f.

2 Vgl. z. B. E. Kreyszig: Statistische Methoden und ihre Anwendungen, 5. erweiterte Auflage, Göttingen 1975, S. 420.

Die eigentliche Bestimmung des Mindeststichprobenumfangs geht von dem Konfidenzintervall[1]

$$p_i - z\sqrt{\frac{P_i(1-P_i)}{m}} \leq P_i \leq p_i + z\sqrt{\frac{P_i(1-P_i)}{m}} \qquad (8)$$

aus, in dem z ein vorgegebener Wert der Standardnormalverteilung ist. Soll die Abweichung der Intervallgrenzen von P_i nicht mehr als e betragen, so ergibt sich $m_{i,min}$, der Mindeststichprobenumfang zur Schätzung von P_i, aus

$$z\sqrt{\frac{P_i(1-P_i)}{m}} = e \qquad (9)$$

als

$$m_{i,min} = \frac{z^2 P_i(1-P_i)}{e^2} \quad . \qquad (10)$$

In der Regel unterscheiden sich für verschiedene Verweilzeiten i die zugehörigen Wahrscheinlichkeiten P_i und damit auch die notwendigen Mindestwerte $m_{i,min}$. Da aus derselben Stichprobe alle n+1 Elemente P_i der Anzahlverteilung geschätzt werden sollen, entspricht der erforderliche Mindeststichprobenumfang m_{min}, der allen Verweilzeiten gerecht wird, dem größten Einzelwert $m_{i,min}$. Der liegt für dasjenige P_i vor, welches 0,5 am nächsten kommt, da das Produkt $P_i(1-P_i)$ sein Maximum bei $P_i = 0,5$ hat.

Nun sind die P_i vor der Erhebung meistens weder bekannt, noch liegen Schätzwerte p_i vor. Wenn man keine Vorkenntnisse besitzt, empfiehlt es sich, $P_i = 0,5$ zu setzen, weil der Mindeststichprobenumfang dann auf keinen Fall unterschätzt werden kann[2]. Man erhält damit

---

1 Vgl. dazu T. Yamane: Statistik - Ein einführendes Lehrbuch, Bd. 2, Frankfurt a.M. 1976, S. 619 ff.

2 Vgl. L. Neumeyer, a.a.O., S. 38.

$$m_{min} = \frac{z^2 \cdot 0,25}{e^2} \quad . \qquad (2)$$

Arbeitet man auf einem Signifikanzniveau von 5%, so ist

$$m_{min} = \frac{1,96^2 \cdot 0,25}{e^2} \qquad (3)$$

nur noch eine Funktion der gewünschten Genauigkeit e. Soll das Konfidenzintervall z. B. nicht breiter als 0,1 sein, die Abweichung also nicht mehr als e=0,05 betragen, so ergibt sich ein notwendiger Stichprobenumfang von m_{min}=385.

Es bleibt anzumerken, daß durch die Nichtbeachtung der Endlichkeit der Grundgesamtheit eine zusätzliche Überschätzung von m_{min} erfolgt, die jedoch bei großen Grundgesamtheiten unbedeutend ist[1].

3.4 Ermittlung der Betragsverteilung nach NEUMEYER

Die Summe der Beträge aller Rechnungen der Grundgesamtheit sei B, derjenige Betrag davon, welcher nach einer Verweilzeit von i Perioden liquidiert wird, sei B_i. Die Komponente W_i der Verweilzeitverteilung der Grundgesamtheit ergibt sich damit als

$$W_i = \frac{B_i}{B} \quad , \qquad i=0,\ldots,n. \qquad (1)$$

Zur Schätzung dieser Komponenten liegen zwei Ansätze von NEUMEYER vor[2], die beide von dem folgenden Urnenmodell ausgehen[3]:

Die Grundgesamtheit[4] besteht gedanklich aus B Kugeln. Jede Kugel repräsentiert eine fakturierte DM[5]. B wird demnach durch

---

1 Vgl. T. Yamane, a.a.O., S. 620: Unter Berücksichtigung der Endlichkeit der Grundgesamtheit müßte m_{min} eigentlich als $m_{min} = \frac{z^2 0,25\ M}{e^2(M-1)+z^2 0,25}$ berechnet werden.

2 Vgl. L. Neumeyer, a.a.O., S. 23 ff.

3 Ebenda, S. 33 ff.

4 NEUMEYER bezeichnet die Teilgesamtheit der bereits vorliegenden Rechnungen als Grundgesamtheit. Danach ist die gesamte Grundgesamtheit erhebbar. Insbesondere ist auch B bekannt. Um die Wiedergabe der NEUMEYERschen Ansätze nicht zu verfälschen, soll eine Umbenennung der dort verwendeten Begriffe in die, welche dieser Arbeit zugrundeliegen, unterbleiben. Wesentliche Einschränkungen inhaltlicher Art ergeben sich dadurch nicht.

5 Zur Vereinfachung wird davon ausgegangen, daß alle Rechnungen nur über volle DM lauten.

die Anzahl der DM-Stücke in dem der Untersuchung zugrundeliegenden Zeitraum angegeben; es ist nichts anderes als der gesamte Umsatz dieses Zeitraums.

Auf jeder Kugel ist vermerkt, wie viele Perioden von der Rechnungserstellung bis zur Bezahlung verstreichen, d. h. wie groß ihre Verweilzeit ist. Für die Schätzung von W_i auf Stichprobenbasis schlägt NEUMEYER ein einstufiges und ein dreistufiges Verfahren vor.

Bei der einstufigen Entnahme wird eine Stichprobe so gezogen, daß alle B DM-Stücke die gleiche Chance haben, in die Stichprobe zu gelangen. Da in einer Rechnung mehrere DM-Stücke gleicher Verweilzeit gewissermaßen zusammengeklebt sind, ist vor der Ziehung eine Aufbereitung des Datenmaterials erforderlich, damit alle Elemente unabhängig voneinander entnommen werden können. Dazu werden sämtliche Rechnungsbeträge in einer Liste gemäß Tabelle 2 kumuliert:

Rechnungsnummer	Verweilzeit in Perioden	Rechnungsbetrag = Anzahl der DM-Stücke dieser Rechnung	Kumulierter Rechnungsbetrag = kum. Anzahl der DM-Stücke
1	4	310	310
2	8	400	710
3	9	26	736
.	.	.	.
:	:	:	:
520	5	130	B= 56 200

Tab. 2: Beispiel für ein Schema zur Kumulierung von Rechnungsbeträgen. Nach: L. NEUMEYER: a.a.O., S. 39.

Die Ziehung der einzelnen DM-Stücke erfolgt nun mit Hilfe von zwischen 1 und B gleichverteilten Zufallszahlen. Lautet die Zufallszahl z. B. 716, so wird die 716. Kugel gezogen. Aus

der Liste kann man ablesen, daß dieses DM-Stück zur 3. Rechnung gehört und eine Verweilzeit von 9 Perioden aufweist.

Die Punkt- und Intervallschätzung von W_i erfolgt bei diesem Entnahmeverfahren analog zur Schätzung von P_i der Anzahlverteilung in Abschnitt 3.2. Die Anzahl der insgesamt gezogenen DM-Stücke sei b. Davon seien b_i solche mit einer Verweilzeit von i. Ein erwartungstreuer Schätzwert für W_i ist dann[1]

$$\hat{W}_i = w_i = \frac{b_i}{b} \quad , \qquad i=0,\ldots,n. \tag{3}$$

Die Varianz von $\hat{W}_i$ beträgt[2]

$$V(\hat{W}_i) = \frac{W_i(1-W_i)}{b} \frac{B-b}{B-1} \tag{4}$$

oder bei großen Grundgesamtheiten, d.h. bei großem Gesamtbetrag B

$$V(\hat{W}_i) = \frac{W_i(1-W_i)}{b} \quad . \tag{5}$$

Sie kann aus der Stichprobe als

$$\hat{V}(\hat{W}_i) = \frac{w_i(1-w_i)}{b-1} \tag{6}$$

geschätzt werden, und das Konfidenzintervall für W_i läßt sich damit als

$$w_i - t\sqrt{\frac{w_i(1-w_i)}{b-1}} \leq W_i \leq w_i + t\sqrt{\frac{w_i(1-w_i)}{b-1}} \quad ,i=0,\ldots,n \tag{7}$$

angeben.

So sehr dieser Ansatz auch durch seine Einfachheit besticht, für die Praxis ist er unbrauchbar, und ein Blick auf Tabelle

---

1 Vgl. L. Neumeyer, a.a.O., S. 36 f. Dieses Verfahren unterstellt ein volldeckendes Spektrum. Unvollständige Bezahlung einer Rechnung läßt sich aber dadurch berücksichtigen, daß man statt des vollen Rechnungsbetrages nur den tatsächlich gezahlten Betrag einsetzt und die Differenz als neue Rechnung mit der Verweilzeit ∞ aufführt.

2 Vgl. L. Neumeyer, a.a.O., S. 36 f.

2 zeigt auch sehr schnell warum. Sämtliche Rechnungen müssen zunächst einmal in einer Liste erfaßt werden, ehe überhaupt eine Stichprobe gezogen werden kann. Wenn aber schon eine Liste erstellt wird, so ist es viel zweckmäßiger, anstelle der fortlaufenden Kumulierung sofort eine nach Verweilzeiten getrennte Kumulierung vorzunehmen und die Verweilzeitverteilung unmittelbar aus allen Fakturen zu berechnen. Die Ziehung der Zufallsstichprobe kann dann unterbleiben. Der einstufige Stichprobenansatz muß sich somit den Vorwurf gefallen lassen, bei zufallsbedingt ungenaueren Ergebnissen aufwendiger zu sein als eine Vollerhebung. Die Ursache dafür ist, daß es so großer Vorbereitungen bedarf, um alle DM-Stücke mit gleicher Wahrscheinlichkeit in die Stichprobe gelangen zu lassen.

Der andere NEUMEYERsche Stichprobenansatz ist als Versuch zu werten, unter Aufrechterhaltung der Unabhängigkeit der Entnahme einzelner DM-Stücke den Aufwand der Vorbereitungen zu reduzieren. Es wird davon ausgegangen, daß die Daten in einer Kartei organisiert sind, die in Ordnern angelegt ist. Jeder Ordner enthält mehrere Blätter, und auf jedem Blatt sind mehrere Rechnungen erfaßt. Auf der 1. Entnahmestufe werden nun einige Ordner stichprobenartig ausgewählt. Jeder dieser Primäreinheiten werden auf der 2. Stufe als Sekundäreinheiten zufallsgesteuert mehrere Blätter entnommen, auf denen dann wie in Tabelle 2 alle Rechnungsbeträge kumuliert werden. Die Kumulierung erfolgt dabei für jedes in die Stichprobe gelangte Blatt gesondert. Die 3. Stufe besteht schließlich darin, daß entsprechend dem Vorgehen beim einstufigen Verfahren pro Blatt mehrere DM-Stücke erhoben werden. Abbildung 6 soll den Ziehungsmechanismus verdeutlichen.

Auf jeder Stufe wird die Ziehung so vorgenommen, daß jedes Element der Stufe die gleiche Chance besitzt, in die Stichprobe zu gelangen.

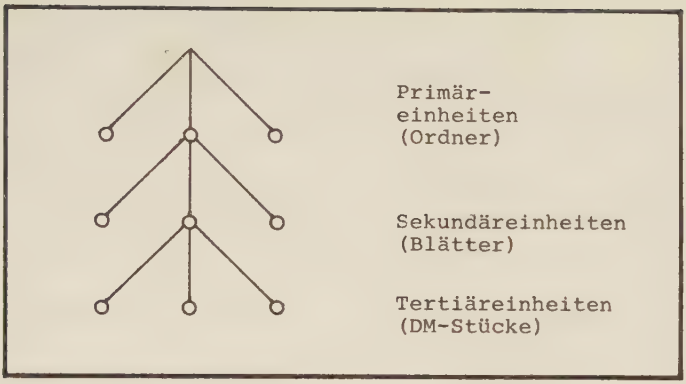

Abb. 6: Entnahmebeispiel für ein dreistufiges Stichprobenverfahren. Nach: L. NEUMEYER, a.a.O., S. 48.

Zur Schätzung von W_i werden beim dreistufigen Verfahren die in Tabelle 3 auf Seite 42 zusammengestellten Größen verwendet.

Die eigentliche Ableitung der Schätzfunktion für W_i und $V(\hat{W}_i)$ ist recht aufwendig und für die Beurteilung des dreistufigen Ansatzes nicht von Bedeutung[1]. Es sollen deshalb nur die Ergebnisse wiedergegeben werden.

NEUMEYER gibt nicht explizit an, welche Funktion er zur Schätzung von W_i verwendet. Aus seinen Ausführungen über die Ermittlung von $\hat{V}(\hat{W}_i)$ geht aber hervor, daß er zu diesem Zweck

$$\hat{W}_i = w_i = \frac{\sum_{j=1}^{m} B_j \, w_{i/j}}{B} \qquad (8)$$

mit

1 Vgl. dazu W. E. Deming: Some Theory of Sampling, New York/London 1950, S. 156; L. Neumeyer, a.a.O., S. 39 ff.

Größe	Bezeichnung in der Grundgesamtheit	Bezeichnung in der Stichprobe
Anzahl der Primäreinheiten (Ordner)	M	m
Anzahl der Sekundäreinheiten (Blätter) in der j-ten Primäreinheit	M_j	m_j
Anzahl der Tertiäreinheiten (DM-Stücke) in der k-ten Sekundäreinheit der j-ten Primäreinheit	B_{jk}	b_{jk}
Mittlere Anzahl der Tertiäreinheiten in einer Sekundäreinheit der j-ten Primäreinheit	$\overline{B}_j = \frac{1}{M_j} \sum_{k=1}^{M_j} B_{jk}$	.
Anzahl der Tertiäreinheiten in der j-ten Primäreinheit	$B_j = \sum_{k=1}^{M_j} B_{jk}$	$b_j = \sum_{k=1}^{m_j} b_{jk}$
Mittlere Anzahl der Tertiäreinheiten in einer Primäreinheit	$\overline{B} = \frac{1}{M} \sum_{j=1}^{M} B_j$	.
Gesamtzahl der Tertiäreinheiten	$B = \sum_{j=1}^{M} B_j$	$b = \sum_{j=1}^{m} b_j$
Relative Häufigkeit des Auftretens eines DM-Stückes mit einer Verweilzeit i	W_i	w_i
Relative Häufigkeit des Auftretens eines DM-Stückes mit einer Verweilzeit i in der j-ten Primäreinheit	$W_{i/j}$	$w_{i/j}$
Relative Häufigkeit des Auftretens eines DM-Stückes mit einer Verweilzeit i in der k-ten Sekundäreinheit der j-ten Primäreinheit	$W_{i/jk}$	$w_{i/jk}$

Tab. 3: Symbole der Größen der dreistufigen Stichprobenerhebung

$$w_{i/j} = \frac{\sum_{k=1}^{m_j} B_{jk} w_{i/jk}}{B_j} \qquad (9)$$

verwendet: Auf jedem Blatt, welches in die Stichprobe gelangt, wird $w_{i/jk}$ gemessen. Dann wird für jeden Ordner das gewogene Mittel $w_{i/j}$ berechnet, wobei die Gesamtbeträge B_{jk} der Blätter als Gewichte dienen. Schließlich wird aus den Mittelwerten pro Blatt ein mit dem Gesamtbetrag der Ordner gewogenes Gesamtmittel w_i festgestellt und zur Schätzung verwendet.

Den Schätzwert für die Varianz $\hat{V}(\hat{W}_i)$ gibt NEUMEYER als[1]

$$\hat{V}(\hat{W}_i) = \frac{1}{m\bar{B}^2} \left\{ \frac{M - m}{M - 1} \hat{\sigma}^2_{i/e} + \frac{1}{M} \sum_{j=1}^{M} \left(\frac{M_j}{m_j}\right)^2 \left[\frac{M_j - m_j}{M_j - 1} m_j \hat{\sigma}^2_{i/j} \right.\right.$$

$$\left.\left. + \frac{m_j}{M_j} \sum_{k=1}^{M_j} \left(\frac{B_{jk}}{b_{jk}}\right)^2 \frac{B_{jk} - b_{jk}}{B_{jk} - 1} b_{jk} \hat{\sigma}^2_{i/jk} \right] \right\} \qquad (10)$$

an. Dabei ist

$\hat{\sigma}^2_{i/jk}$ ein Schätzwert für die Streuung der Ereignisfälle "DM-Stück hat die Verweilzeit i" innerhalb der k-ten Sekundäreinheit der j-ten Primäreinheit,

$\hat{\sigma}^2_{i/j}$ ein Schätzwert für die Streuung dieser Ereignisfälle innerhalb der j-ten Primäreinheit,

$\hat{\sigma}^2_{i/e}$ ein Schätzwert für die äußere Streuung der Ereignisfälle je Primäreinheit.

Die Schätzwerte berechnen sich als

$$\hat{\sigma}^2_{i/jk} = \frac{B_{jk} - 1}{B_{jk}} \cdot \frac{b_{jk}}{b_{jk} - 1} s^2_{i/jk} \text{ mit } s^2_{i/jk} = w_{i/jk}(1 - w_{i/jk}), \qquad (11)$$

$$\hat{\sigma}^2_{i/j} = \frac{M_j - 1}{m_j - 1} \left(\frac{m_j}{m_j - 1} s^2_{i/j} - \frac{1}{m_j} \sum_{k=1}^{m_j} B^2_{jk} \frac{\hat{\sigma}^2_{i/jk}}{b_{jk}} \cdot \frac{B_{jk} - b_{jk}}{B_{jk} - 1} \right) \qquad (12)$$

---

1 Zur Ableitung von $\hat{V}(\hat{W}_i)$ vgl. W. E. Deming, a.a.O., S. 142 ff; L. Neumeyer, a.a.O., S. 44 ff.

mit
$$s^2_{i/j} = \frac{\sum_{k=1}^{m_j} (B_{jk}\, w_{i/jk} - \bar{B}\, w_{i/j})^2}{m_j} , \tag{13}$$

$$\hat{\sigma}^2_{i/e} = \frac{M-1}{m} \frac{m}{m-1} s^2_{i/e} - \frac{M-1}{m\,M} \sum_{j=1}^{m} M_j^2 \frac{\hat{\sigma}^2_{i/j}}{m_j} \frac{M_j - m_j}{m_j - 1} \tag{14}$$

mit
$$s^2_{i/e} = \frac{\sum_{j=1}^{m} (B_j\, w_{i/j} - \bar{B}\, w_i)^2}{m} . \tag{15}$$

Das Konfidenzintervall lautet schließlich

$$w_i - t\sqrt{\hat{V}(\hat{W}_i)} \leq W_i \leq w_i + t\sqrt{\hat{V}(\hat{W}_i)} .$$

Dieser Ansatz ist im Gegensatz zur oben beschriebenen Auswahl ein wirkliches Stichprobenverfahren, da er die vollständige Kenntnis der Grundgesamtheit nicht voraussetzt.

Eine eindeutige Bestimmung des Stichprobenumfangs, der eine vorgegebene Intervallbreite gewährleistet, ist hier nicht möglich, weil durch die Aufteilung der Tertiäreinheiten auf die Primär- und Sekundäreinheiten zwei Freiheitsgrade entstehen. NEUMEYER schlägt vor, 2 Ordner und aus diesen je 10 Blätter auszuwählen, denen dann jeweils 20 DM-Stücke entnommen werden sollen[1]. Die Kumulierung der Rechnungen hat also nur noch für 20 Blätter zu erfolgen. Der Arbeitsaufwand der Erhebung wird dadurch gegenüber dem einstufigen Verfahren drastisch gesenkt.

Im Hinblick auf die Genauigkeit der Schätzung von W_i ist bei einer vorgegebenen Anzahl von 400 Tertiäreinheiten die Auswahl von $m = 2$ Primäreinheiten und daraus jeweils die von $m_j = 10$ Sekundäreinheiten aber denkbar ungünstig. Gleichung (10) macht

---

1 Vgl. L. Neumeyer, a.a.O., S. 52.

das deutlich[1]: Während sämtliche Summanden innerhalb der geschweiften Klammer durch m dividiert werden, hat der Betrag von m_j nur einen Einfluß auf die eckige Klammer. Es wäre deshalb bei gegebener Anzahl von Tertiäreinheiten günstiger, m zu Lasten von m_j zu vergrößern.

Darüberhinaus müssen auch erhebliche Genauigkeitseinbußen bei der Intervallschätzung von W_i hingenommen werden. In der Schätzfunktion (10) nimmt $\hat{\sigma}^2_{i/e}$ einen wesentlichen Einfluß auf $\hat{V}(\hat{W}_i)$. Die Berechnung von $\hat{\sigma}^2_{i/e}$ erfolgt nach (14) in Verbindung mit (15). Wählt man nun m = 2, so wird die Varianz $\hat{\sigma}^2_{i/e}$ - von der Korrektur in (14) abgesehen - aus nur zwei Werten geschätzt. Dieser Schätzung wird man kein großes Vertrauen entgegenbringen.

Um eine höhere Genauigkeit sowohl der Punkt- als auch der Intervallschätzung von W_i zu erzielen, müßte man die Anzahl der Primäreinheiten vergrößern. Gleichzeitig dürfte die Anzahl der Sekundäreinheiten pro Primäreinheit nicht zu klein werden, da die Ungenauigkeit sonst lediglich verlagert würde (z. B. auf die Schätzung von $\hat{\sigma}^2_{i/j}$). Insgesamt müßten also mehr Sekundäreinheiten einbezogen werden, demnach auch mehr Listen erstellt werden, und damit würde der Erhebungsaufwand sehr schnell steigen.

3.5 Ermittlung der Betragsverteilung durch Verhältnisschätzung

Im Kern der NEUMEYERschen Ansätze steht das Bestreben, einzelne DM-Stücke voneinander unabhängig zu ziehen, um Verzerrungen bei der Schätzung von W_i auszuschalten. Die Erwartungstreue wird dabei jedoch mit einem so großen Aufwand bei der Erhebung erkauft, daß sich die Frage stellt, ob nicht unter Verzicht auf

---

1 Vgl. dazu auch H. Kellerer: Theorie und Technik des Stichprobenverfahrens - Eine Einführung unter besonderer Berücksichtigung der Anwendung auf soziale und wirtschaftliche Massenerscheinungen, Einzelschriften der Deutschen Statistischen Gesellschaft, 3. Aufl., München 1963, S. 148.

die Unverzerrtheit ein praktikableres Stichprobenverfahren gewählt werden sollte. Erheblich einfacher als die Entnahme einzelner DM-Stücke ist die Ziehung ganzer Rechnungen durchzuführen, die dem im folgenden vorgestellten Verfahren zugrunde liegt.

Dieses Stichprobenverfahren unterstellt dasselbe Urnenmodell, wie es NEUMEYER verwendet. Jede fakturierte DM stellt gedanklich also wieder eine Kugel dar, auf der die Verweilzeit vermerkt ist, und mehrere Kugeln gleicher Verweilzeit sind wieder zu einem Klumpen (einer Rechnung) zusammengeklebt. Die Ziehung erfolgt nun jedoch dergestalt, daß alle Klumpen mit der gleichen Wahrscheinlichkeit in die Stichprobe gelangen können.

Auf diese Weise werden m Rechnungen entnommen. Die j-te Rechnung in der Stichprobe laute über einen Betrag von r_j DM. Davon seien b_{ij} DM-Stücke solche mit einer Verweilzeit i. Werden Teil- und Überzahlung ausgeschlossen, so ist

$b_{ij} = 0$, wenn die Verweilzeit der Rechnung ungleich i ist,

$b_{ij} = r_j$, wenn die Verweilzeit i ist und keine Einnahmeminderungen, etwa durch Inanspruchnahme von Skonto, auftreten,

$0 < b_{ij} < r_j$, wenn die Verweilzeit i ist und Einnahmeminderungen auftreten.

Da $$W_i = \frac{B_i}{B} = \frac{\sum_{j=1}^{M} b_{ij}}{\sum_{j=1}^{M} r_j}, \qquad (1)$$

liegt es nahe, als Schätzwert für W_i

$$\hat{W}_i = w_i = \frac{\sum_{j=1}^{m} b_{ij}}{\sum_{j=1}^{m} r_j} \qquad (2)$$

zu verwenden[1]. Dabei ist $\sum_{j=1}^{m} b_{ij}$ der insgesamt in der Stichprobe beobachtete Betrag mit einer Verweilzeit von i und $\sum_{j=1}^{m} r_j$ der gesamte Stichprobenrechnungsbetrag. Erweitert man (2) mit $\frac{1}{m}$, so ergibt sich

$$w_i = \frac{\frac{1}{m} \sum_{j=1}^{m} b_{ij}}{\frac{1}{m} \sum_{j=1}^{m} r_j}, \tag{3}$$

oder mit

$$\bar{b}_i = \frac{1}{m} \sum_{j=1}^{m} b_{ij} \tag{4}$$

und

$$\bar{r} = \frac{1}{m} \sum_{j=1}^{m} r_j, \tag{5}$$

$$w_i = \frac{\bar{b}_i}{\bar{r}}. \tag{6}$$

Der Schätzwert für W_i ist demnach der Quotient aus der durchschnittlichen Anzahl von DM-Stücken mit einer Verweilzeit von i pro Rechnung und dem mittleren Rechnungsbetrag in der Stichprobe.

Die Stichprobenwerte $\bar{b}_i$ und $\bar{r}$ sind erwartungstreue Schätzwerte für die entsprechenden Parameter der Grundgesamtheit, die als

$$\bar{B}_i = \frac{1}{M} \sum_{j=1}^{M} b_{ij} \tag{7}$$

und

$$\bar{R} = \frac{1}{M} \sum_{j=1}^{M} r_j \tag{8}$$

---

1 Vgl. H. Langen et al., a.a.O., S. 53.
EDIN erhält dieselbe Schätzfunktion mit Hilfe der Maximum-likelihood-Methode (vgl. R. Edin, Dynamische Analyse ..., a.a.O., S. 167 ff.). Er geht dabei jedoch (ohne es explizit anzugeben) von der sehr wirklichkeitsfremden Annahme aus, daß die einzelnen DM-Stücke einer Rechnung voneinander unabhängig bezahlt werden.

definiert sind. Für ihren Quotienten gilt das jedoch nicht. $\hat{W}_i$ ist zwar eine konsistente, aber keine unverzerrte Schätzung[1] für W_i. Das Ausmaß der Verzerrung läßt sich nicht e x a k t angeben. Es hat die Größenordnung $\frac{1}{m}$ und ist damit bei großem Stichprobenumfang m unbedeutend[2]. Dann gilt die Näherungsgleichung

$$E(\hat{W}_i) \doteq W_i \quad . \tag{9}$$

Auch für die Varianz $V(\hat{W}_i)$ existiert keine exakte Formel. Es läßt sich aber zeigen, daß bei großem m für $V(\hat{W}_i)$ näherungsweise[3]

$$V(\hat{W}_i) \doteq \frac{1}{m\,\bar{R}^2} \; \frac{\sum_{j=1}^{M} (b_{ij}-W_i r_j)^2}{M} \; \frac{M-m}{M-1} \tag{10}$$

gilt. Aufgrund der Annahme einer im Vergleich zum Stichprobenumfang sehr großen Grundgesamtheit vereinfacht sich (10) zu

$$V(\hat{W}_i) \doteq \frac{1}{m\,\bar{R}^2} \; \frac{\sum_{j=1}^{M} (b_{ij}-W_i r_j)^2}{M} \quad . \tag{11}$$

Faßt man b_{ij} als j-te Realisation einer Zufallsvariablen[4] B_i auf und r_j als j-te Realisation einer Zufallsvariablen R, so läßt sich $V(\hat{W}_i)$ unter Verwendung der Varianzen $V(B_i)$ und $V(R)$ und der Kovarianz $C(B_i,R)$ gleichwertig als

$$V(\hat{W}_i) \doteq \frac{1}{m\,\bar{R}^2} \left[V(B_i) - 2\,W_i\,C(B_i,R) + W_i^2 V(R) \right] \tag{12}$$

---

1 Vgl. z. B. W. G. Cochran, a.a.O., S. 46 ff; W. E. Deming, a.a.O., S. 165 ff; M. H. Hansen, W. N. Hurwitz u. W. G. Madow: Sample Survey Methods and Theory, Vol. 1, Methods and Applications, New York/London/Sydney 1953, S. 36 ff. u. S. 126 ff.

2 Vgl. W. G. Cochran, a.a.O., S. 47 f.

3 Ebenda.

4 Um Verwechslungen mit den Parametern der Grundgesamtheit zu vermeiden, die ebenfalls mit Großbuchstaben bezeichnet werden, sollen einige Zufallsvariablen *kursiv* geschrieben werden.

ausdrücken[1]. Diese Schreibweise ist zwar die geläufigere, für den weiteren Gang der Untersuchung erweist sich jedoch (11) als vorteilhafter.

$\hat{V}(\hat{W}_i)$ wird nun aus der Stichprobe als

$$\hat{V}(\hat{W}_i) = \frac{1}{m\,\bar{r}^2} \sum_{j=1}^{m} \frac{(b_{ij}-w_i r_j)^2}{m-1} \qquad (13)$$

geschätzt. Zur einfacheren Berechnung empfiehlt sich noch eine Umformung von (13) zu

$$\hat{V}(\hat{W}_i) = \frac{1}{m\,\bar{r}^2(m-1)} \left(\sum_{j=1}^{m} b_{ij}^2 - 2w_i \sum_{j=1}^{m} b_{ij} r_j + w_i^2 \sum_{j=1}^{m} r_j^2 \right). \qquad (14)$$

Ist m genügend groß, so ist $\hat{W}_i$ annähernd normalverteilt[2]. Dann läßt sich für W_i die Intervallschätzung

$$w_i - t\sqrt{\hat{V}(\hat{W}_i)} \leq W_i \leq w_i + t\sqrt{\hat{V}(\hat{W}_i)} \qquad (15)$$

abgeben. Den Wert von t entnimmt man dabei der entsprechenden t-Verteilung mit m-1 Freiheitsgraden oder man verwendet den bereits oben angegebenen Faustwert $t=2$, sofern ein Signifikanzniveau von $\alpha=0{,}05$ gewünscht wird.

Was unter genügend groß zu verstehen ist, kann nur sehr grob angegeben werden. "Die Verteilung von Verhältnisschätzwerten ist ausgesprochen schlecht zugänglich... Das bekannte theoretische Wissen ist so schlecht, daß es für Anwendungsfälle kaum brauchbar ist."[3] Als grobe Faustregel gibt COCHRAN an, daß die Approximation durch die Normalverteilung verwendet werden darf, wenn der Stichprobenumfang mindestens m=30 beträgt und darüber-

---

1 Vgl. W. G. Cochran, a.a.O., S. 190.

2 Ebenda, S. 47.

3 Ebenda, S. 188.

hinaus so groß ist, daß die Variationskoeffizienten von $\bar{b}_i$ und $\bar{r}$ jeweils kleiner als 0,1 sind[1]. Diese Koeffizienten können aus der Stichprobe als

$$\hat{CV}_{\bar{b}_i} = \frac{\sqrt{\hat{V}(\bar{b}_i)}}{\bar{b}_i} = \frac{\sqrt{\frac{1}{m}\sum_{j=1}^{m}\frac{(b_{ij}-\bar{b}_i)^2}{m-1}}}{\bar{b}_i} \tag{16}$$

und

$$\hat{CV}_{\bar{r}} = \frac{\sqrt{\hat{V}(\bar{r})}}{\bar{r}} = \frac{\sqrt{\frac{1}{m}\sum_{j=1}^{m}\frac{(r_j-\bar{r})^2}{m-1}}}{\bar{r}} \tag{17}$$

geschätzt werden.

3.6 Bestimmung des Mindeststichprobenumfangs zur Ermittlung der Betragsverteilung durch Verhältnisschätzung

Die Bestimmung des notwendigen Stichprobenumfangs bei der Schätzung der Betragsverteilung geht von dem Konfidenzintervall der Verhältnisschätzung

$$w_i - z\sqrt{V(\hat{W}_i)} \leq W_i \leq w_i + z\sqrt{V(\hat{W}_i)} \tag{1}$$

aus. Den für eine gewünschte Intervallbreite von 2e bei einem vorgegebenen Signifikanzniveau (ausgedrückt duch z) erforderlichen Mindeststichprobenumfang zur Schätzung von W_i erhält man bei großen Grundgesamtheiten aus

$$z\sqrt{\frac{1}{m\bar{R}^2}\sum_{j=1}^{M}\frac{(b_{ij}-W_i r_j)^2}{M}} = e \tag{2}$$

als

$$m_{i,min} = \frac{z^2\frac{1}{\bar{R}^2}\sum_{j=1}^{M}\frac{(b_{ij}-W_i r_j)^2}{M}}{e^2} \; . \tag{3}$$

---

1 Vgl. W. G. Cochran, a.a.O., S. 189.

Wie bei der Ermittlung der Anzahlverteilung werden sich auch nach (3) für die einzelnen Komponenten W_i unterschiedliche Werte für $m_{i,min}$ ergeben. Der größte davon ist der gesuchte Mindeststichprobenumfang m_{min}.

Bei der Anzahlbetrachtung läßt sich durch Vorgabe von $P_i=1/2$ ein Stichprobenumfang berechnen, der auf jeden Fall den Genauigkeitsanforderungen entspricht. Dieser Gedanke kann hier nicht unmittelbar übernommen werden. Es sind vielmehr zusätzliche Kenntnisse über die Rechnungsbeträge r_j und die eingegangenen Beträge b_{ij} erforderlich. Die b_{ij} kennzeichnen das Zahlungsverhalten der Kunden, das durch die Stichprobenentnahme erst noch ermittelt werden soll. Es kann deshalb nicht davon ausgegangen werden, daß über sie bereits Kenntnisse vorliegen. Dagegen ist es durchaus denkbar, daß man aus anderen Untersuchungen schon eine Vorstellung über den Mittelwert und die Varianz der Rechnungsbeträge r_j gewonnen hat, und in diesem Fall läßt sich zumindest die Größenordnung von m_{min} abschätzen.

Dazu sei der Spezialfall betrachtet, daß Unabhängigkeit zwischen dem fakturierten Betrag einer Rechnung und ihrer Verweilzeit besteht. Dann existiert eine feste Beziehung zwischen der Anzahl- und der Betragsverteilung in der Grundgesamtheit, die zunächst aufgezeigt werden soll.

Laut Definition ist

$$P_i = \frac{A_i}{M} , \qquad (4)$$

wobei A_i die Anzahl der Rechnungen ist, die nach i Perioden beglichen werden, und M ihre gesamte Anzahl in der Grundgesamtheit. Führt man eine Zufallsvariable X_i ein, welche die Ausprägung $x_{ij} = 1$ annimmt, wenn die j-te Rechnung eine Verweilzeit von i Perioden hat, und sonst den Wert $x_{ij} = 0$, so läßt sich aus P_i auch als

$$P_i = \frac{1}{M} \sum_{j=1}^{M} x_{ij} = E(\chi_i) \qquad (5)$$

schreiben. Wie immer bei der Verwendung der Anzahlverteilung wird auch hier Teilzahlung ausgeschlossen.

Die Komponenten der Betragsverteilung sind definiert als

$$W_i = \frac{B_i}{B} \qquad (6)$$

mit

$$B_i = \sum_{j=1}^{M} b_{ij} \qquad (7)$$

und

$$B = \sum_{j=1}^{M} r_j \; . \qquad (8)$$

Der Einfachheit halber sei unterstellt, daß alle Kunden bei der Bezahlung einer Rechnung nach i Perioden einheitlich denselben Skontosatz c_i in Anspruch nehmen[1].

Dann läßt sich b_{ij} auch als

$$b_{ij} = r_j(1-c_i)x_{ij} \qquad (9)$$

ausdrücken, wobei x_{ij} die Ausprägung der oben eingeführten Zufallsvariablen χ_i ist. Damit wird (6) zu

$$W_i = \frac{\sum_{j=1}^{M} r_j(1-c_i)x_{ij}}{B} = (1-c_i)\,\frac{\frac{1}{M}\sum_{j=1}^{M} r_j x_{ij}}{\frac{1}{M}\,B} \qquad (10)$$

und mit

$$\bar{R} = \frac{1}{M}\,B$$

---

1 Diese Vereinfachung erscheint vertretbar, da solche Einnahmeminderungen meist nur geringe Bruchteile der Gesamteinnahmen betragen und die Berücksichtigung individueller Betragskürzungen in diesem Zusammenhang zu keinen nennenswerten Abweichungen im Ergebnis führen. Der vollständige Ausfall einer Forderung wird durch die Annahme nicht berührt. Er wird dadurch berücksichtigt, daß der entsprechenden Rechnung keine bzw. eine gedachte sehr große Verweilzeit zugeordnet wird.

zu

$$W_i = (1-c_i) \frac{\frac{1}{M} \sum_{j=1}^{M} r_j x_{ij}}{\bar{R}} \tag{11}$$

Faßt man r_j wieder als Realisation einer Zufallsvariablen R mit dem Erwartungswert $\bar{R}$ auf, so ergibt sich unter Verwendung des Erwartungswertoperators

$$W_i = (1-c_i) \frac{E(RX_i)}{\bar{R}} . \tag{12}$$

Bei Unabhängigkeit von Rechnungsbeträgen und Verweilzeiten gilt

$$W_i = (1-c_i) \frac{E(R)\ E(X_i)}{\bar{R}} = (1-c_i)P_i . \tag{13}$$

Die Betragsverteilung stellt dann die um die Einnahmeminderungen korrigierte Anzahlverteilung dar. Treten keine Betragskürzungen auf, so ist $c_i = 0$ für alle i, und die Anzahlverteilung ist identisch mit der Betragsverteilung.

Unter der Annahme der Unabhängigkeit soll nun der Mindeststichprobenumfang $m_{i,min}$ aus

$$z \sqrt{\hat{V}(\hat{W}_i)} = e \tag{14}$$

ermittelt werden.

$$\hat{V}(\hat{W}_i) \doteq \frac{1}{m \bar{R}^2} \sum_{j=1}^{M} \frac{(b_{ij}-W_i r_j)^2}{M} \tag{15}$$

läßt sich umformen zu

$$\hat{V}(\hat{W}_i) \doteq \frac{1}{m \bar{R}^2} \frac{1}{M} \sum_{j=1}^{M} r_j^2 \left(\frac{b_{ij}}{r_j} - W_i\right)^2 . \tag{16}$$

Aus (9) erhält man

$$\frac{b_{ij}}{r_j} = (1-c_i)x_{ij} . \tag{17}$$

Damit ist

$$V(\hat{W}_i) \doteq \frac{1}{m\,\bar{R}^2}\ \frac{1}{M}\sum_{j=1}^{M} r_j^2\left[(1-c_i)x_{ij}-W_i\right]^2 \qquad (18)$$

oder, bei Verwendung des Erwartungswertoperators,

$$V(\hat{W}_i) \doteq \frac{1}{m\,\bar{R}^2}\ E\left[R^2((1-c_i)X_i-W_i)^2\right]. \qquad (19)$$

Wegen der Unabhängigkeit zwischen Betrag und Verweilzeit besteht auch Unabhängigkeit zwischen R und X_i. Aus (19) ergibt sich deshalb

$$\begin{aligned} V(\hat{W}_i) &\doteq \frac{1}{m\,\bar{R}^2}\ E(R^2)E\left[(1-c_i)X_i-W_i\right]^2 \\ &\doteq \frac{1}{m\,\bar{R}^2}\ E(R^2)E\left[(1-c_i)^2X_i^2-2(1-c_i)W_iX_i+W_i^2\right] \\ &\doteq \frac{1}{m\,\bar{R}^2}\ E(R^2)\left[(1-c_i)^2E(X_i^2)-2(1-c_i)W_iE(X_i)+W_i^2\right]. \end{aligned} \qquad (20)$$

Mit

$$E(X_i) = E(X_i^2) = P_i \qquad (21)$$

läßt sich (20) vereinfachen zu

$$V(\hat{W}_i) \doteq \frac{1}{m\,\bar{R}^2}\ E(R^2)\left[(1-c_i)^2P_i-2(1-c_i)W_iP_i+W_i^2\right] \qquad (22)$$

und unter Berücksichtigung von $W_i=(1-c_i)P_i$ ergibt sich

$$V(\hat{W}_i) \doteq \frac{1}{m\,\bar{R}^2}\ E(R^2)\left[(1-c_i)^2P_i(1-P_i)\right]. \qquad (23)$$

In Verbindung mit (14) erhält man den gesuchten Mindeststichprobenumfang als

$$m_{i,min} = \frac{z^2\,\frac{E(R^2)}{\bar{R}^2}\left[(1-c_i)^2P_i(1-P_i)\right]}{e^2}. \qquad (24)$$

Bei festem $E(R^2)$, $\bar{R}$ und c_i hat $m_{i,min}$ sein Maximum bei $P_i=0,5$. Der Mindeststichprobenumfang zur Ermittlung der Betragsvertei-

lung lautet damit bei Unabhängigkeit von Betrag und Verweilzeit

$$m_{min} = \frac{z^2 \frac{E(R^2)}{\bar{R}^2} (1-c_i)^2 \cdot 0{,}25}{e^2} \qquad (25)$$

oder wegen

$$V(R) = E(R^2) - \bar{R}^2$$

$$m_{min} = \frac{0{,}25z^2 \frac{V(R)+\bar{R}^2}{\bar{R}^2}(1-c_i)^2}{e^2} = \frac{0{,}25z^2}{e^2}\left[1 + \frac{V(R)}{\bar{R}^2}\right](1-c_i)^2. \qquad (26)$$

Der notwendige Stichprobenumfang ist also außer von z und e auch noch vom Quadrat des Variationskoeffizienten $\sqrt{V(R)}/\bar{R}$ der Rechnungsbeträge abhängig. Für $V(R)=0$ entspricht er gerade dem Mindeststichprobenumfang bei der Schätzung der Anzahlverteilung. Dieser Fall ist aber sehr unrealistisch, weil sämtliche Rechnungen den gleichen Betrag aufweisen müßten. Ist $V(R)>0$, so ist der Stichprobenumfang größer als bei der Anzahlschätzung.

Die Gleichung (26) gilt nur bei Unabhängigkeit von Beträgen und Verweilzeiten. Das ist dann der Fall, wenn nicht mit bestimmten Werten von r_j bestimmte Werte von i besonders häufig auftreten. Insbesondere bei einer großen Anzahl von Kunden können sich die Einflüsse des individuellen Zahlungsverhaltens auf die Verweilzeitverteilung aufheben, so daß insgesamt Unabhängigkeit resultieren kann; es braucht aber nicht immer so zu sein. Welcher Zustand vorliegt, ist vor der Erhebung nicht bekannt. Die Aussagekraft des nach (26) berechneten Mindeststichprobenumfangs wird dadurch zwar eingeschränkt, in Ermangelung anderer Anhaltspunkte kann m_{min} dennoch, auch wenn keine Unabhängigkeit gegeben ist, zumindest zur Angabe der Größenordnung der Stichprobe dienen.

Wenn $V(R)/\bar{R}^2$ vor der Stichprobenentnahme unbekannt ist, läßt sich a-priori nicht einmal mehr die Größenordnung von m_{min} abschätzen.

Man kann dann lediglich mit $0{,}25z^2/e^2$ eine untere Grenze[1] für m_{min} angeben. In diesem Fall empfiehlt sich ein zweistufiges Vorgehen:

1. Man zieht $0{,}25z^2/e^2$ Rechnungen, schätzt $m_{i,min}$ $(i=0,\ldots,n)$ als

$$\hat{m}_{i,min} = \frac{z^2 \frac{1}{\bar{r}^2} \sum_{j=1}^{m} \frac{(b_{ij}-w_i r_j)^2}{m-1}}{e^2} \qquad (27)$$

   und ermittelt $\hat{m}_{min}$ als Maximum der $\hat{m}_{i,min}$. Die Berechnung der $\hat{m}_{i,min}$ bedeutet dabei keinen wesentlichen Zusatzaufwand, weil die dazu verwendeten Größen ohnehin zur Schätzung von $V(\hat{W}_i)$ benötigt werden.

2. Man entnimmt die noch an $\hat{m}_{min}$ fehlenden Rechnungen.

Es sei noch darauf hingewiesen, daß für den Sonderfall der Unabhängigkeit von Betrag und Verweilzeit ein unmittelbarer Vergleich der Schätzgenauigkeiten bei der Ermittlung von Anzahl- und Betragsverteilung möglich ist. Die Varianz $V(\hat{W}_i)$ lautet dann

$$V(\hat{W}_i) \doteq \frac{1}{m} P_i(1-P_i)(1-c_i)^2\left(1+\frac{V(R)}{\bar{R}^2}\right) , \qquad (28)$$

und der Vergleich mit

$$V(\hat{P}_i) = \frac{1}{m} P_i(1-P_i) \qquad (29)$$

zeigt, daß die Genauigkeit der Schätzung $\hat{W}_i$ gegenüber $\hat{P}_i$ bei festem m mit zunehmendem Variationskoeffizienten $\sqrt{V(R)}/\bar{R}$ rasch abnimmt. (Der Faktor $(1-c_i)^2$ spielt wegen der geringen Größe von c_i nur eine untergeordnete Rolle.) Wenn auch der genaue funktionale Zusammenhang zwischen $V(\hat{W}_i)$ und $\sqrt{V(R)}/\bar{R}$ nur unter der obigen Annahme besteht, so kann doch für den allgemeinen Fall zumindest größenordnungsmäßig eine ähnliche Beziehung unterstellt werden.

---

1 Den Faktor $(1-c_i)^2$ wird man bei der praktischen Anwendung häufig unberücksichtigt lassen können.

3.7 Modifizierte Ermittlung der Verweilzeitverteilung

3.7.1 Ermittlung aus einer vollständigen Stichprobe

Alle vorgestellten Verfahren zur Ermittlung von Verweilzeitverteilungen gehen davon aus, daß man von jeder fakturierten DM bzw. jeder Rechnung in der Stichprobe die Verweilzeit feststellen kann. Beträgt die längste Verweilzeit n Perioden und sind alle Fakturen, die in die Stichprobe gelangen können, mindestens n Perioden alt, so ist dieses Urnenmodell angemessen. Man kann dann für jede gezogene Rechnung die Verweilzeit angeben, da ihre Bezahlung - sofern sie überhaupt erfolgt - bis zur Periode der Stichprobenziehung stattgefunden hat. Rechnungen, die jünger als n Perioden sind, können in dem verwendeten Urnenmodell nicht berücksichtigt werden, selbst dann nicht, wenn sie zum Erhebungszeitpunkt schon beglichen worden sind, ihre Verweilzeit also bekannt ist. Jede Rechnung, die in die Ermittlung der Verweilzeitverteilung eingehen soll, muß nämlich jede der möglichen Verweilzeiten i=0,...,n aufweisen können. Diese Forderung wird aber nicht von denjenigen Fakturen erfüllt, deren Erstellung weniger als n Perioden zurückliegt.

Bei langen Verweilzeitverteilungen, also bei großem n, ergibt sich die ungünstige Situation, daß das zur Ermittlung verwendete Datenmaterial schon bei der Erstellung der ersten Prognose recht alt ist. Da ferner die längste Verweilzeit n vor der Erhebung nicht bekannt ist, läßt sich nicht einmal angeben, aus welchem Zeitraum die Daten für die Stichprobenentnahme herangezogen werden sollen. Liegen keine Anhaltspunkte aus früheren Untersuchungen vor und soll das Urnenmodell beibehalten werden, so bleibt nichts anderes übrig, als eine intuitive und zwangsläufig recht grobe Schätzung des relevanten Zeitraums vorzunehmen. Daraus wird dann die Stichprobe gezogen. Man ermittelt die längste beobachtete Verweilzeit n und eliminiert schließlich alle Rechnungen aus der Stichprobe, die älter als n Perioden sind.

Nun ist es ausgesprochen unbefriedigend, bereits gezogene Rechnungen wieder aus der Stichprobe zu entfernen, besonders wenn es sich um diejenigen handelt, welche die neuesten Informationen über das Zahlungsverhalten liefern. Theoretisch wird zwar Stabilität der Verweilzeitverteilung unterstellt, in der Praxis sind jedoch kleinere Abweichungen zu erwarten. Man wird deshalb bestrebt sein, möglichst junge Daten in die Stichprobe gelangen zu lassen, und durch eine leichte Modifikation der Schätzverfahren, welche auf der Entnahme ganzer Rechnungen beruhen[1], kann erreicht werden, daß a l l e gezogenen Fakturen, deren Verweilzeit zum Erhebungszeitpunkt bekannt ist, bei der Ermittlung der Verweilzeitverteilung berücksichtigt werden.

Zur Vereinfachung der Darstellung sei zunächst unterstellt, daß nur die drei Verweilzeiten 0, 1 und 2 existieren. Ferner seien ausschließlich Rechnungen betrachtet, die in den letzten vier Perioden (einschließlich der Erhebungsperiode) erstellt worden sind. Das Ergebnis einer Stichprobe von m Fakturen zur Schätzung der Anzahlverteilung sei in der folgenden Matrix wiedergegeben:

i \ t	t=0	t=-1	t=-2	t=T=-3	$\Sigma_{t=-i}^{-3} a_{it}$
i=0	a_{00}	$a_{0,-1}$	$a_{0,-2}$	$a_{0,-3}$	a'_0
i=1	...	$a_{1,-1}$	$a_{1,-2}$	$a_{1,-3}$	a'_1
i=2	...	...	$a_{2,-2}$	$a_{2,-3}$	a'_2
m_t	m_0	m_{-1}	m_{-2}	m_{-3}	·

Tabelle 4: Matrix der Ergebnisse einer Stichprobe zur Ermittlung der Anzahlverteilung

1 Auch die NEUMEYERschen Ansätze müssen modifiziert werden, wenn sie die jüngeren Rechnungen mitberücksichtigen wollen. Darauf soll aber nicht eingegangen werden.

Dabei ist

t der Periodenindex,

i die Verweilzeit,

a_{it} die Anzahl der Rechnungen in der Stichprobe, die eine Verweilzeit von i Perioden haben und aus der Periode t stammen,

a'_i die Anzahl der Rechnungen in der Stichprobe, bei denen die Verweilzeit i b e o b a c h t e t wird (der Strich dient der Unterscheidung von a_i, der t a t s ä c h l i c h vorhandenen Anzahl mit einer Verweilzeit von i). Es gilt also $a'_i = \sum_{t=-i}^{T} a_{it}$, wobei T der Index der am weitesten zurückliegenden Periode in der Stichprobe ist.

m_t die Anzahl der Rechnungen in der Stichprobe, die in der Periode t erstellt wurden, wobei $m=\sum_t m_t$.

Die links unten in der Tabelle stehenden Felder können nicht ausgefüllt werden, da die entsprechenden Verweilzeiten in der Erhebungsperiode noch nicht beobachtet werden können.

Bei allen m Rechnungen der Stichprobe kann man in der Erhebungsperiode t=0 angeben, ob sie eine Verweilzeit von 0 Perioden haben oder nicht. Das Element P_0 der Anzahlverteilung wird deshalb wie zuvor als

$$\hat{P}_0 = p_0 = \frac{a'_0}{m} \qquad (1)$$

geschätzt.

Die mit der Verweilzeit 1 beobachteten Rechnungen können nur aus den Perioden t=-1, -2 und -3 stammen. Ihre Anzahl a'_1 ist zur Schätzung von P_1 daher nicht mit m, sondern nur mit der Anzahl der aus diesen Perioden stammenden Rechnungen der Stichprobe in Beziehung zu setzen:

$$\hat{P}_1=p_1= \frac{a'_1}{m_{-1}+m_{-2}+m_{-3}} \qquad (2)$$

oder, wegen

$$m = \sum_{t=0}^{-3} m_t ,$$

$$\hat{P}_1 = \frac{a_1'}{m-m_0} . \qquad (3)$$

Der Schätzwert für P_2 lautet entsprechend

$$\hat{P}_2 = p_2 = \frac{a_2'}{m-m_0-m_{-1}} , \qquad (4)$$

und P_i $(i \neq 0)$ läßt sich allgemein als

$$\hat{P}_i = p_i = \frac{a_i'}{m - \sum_{t=0}^{-i+1} m_t} , \quad i=1,\ldots,n, \qquad (5)$$

schätzen.

Die Anzahl der zur Schätzung von P_i herangezogenen Rechnungen beträgt demnach $m-\sum_{t=0}^{-i+1} m_t$. Hätte man das bisher verwendete Urnenmodell beibehalten und alle Fakturen, die jünger als n Perioden sind, aus der Stichprobe eliminiert, so wären sämtliche Elemente P_i nur aus $m-\sum_{t=0}^{-n+1} m_t$ Rechnungen geschätzt worden. Die durch die Berücksichtigung aller bis zum Erhebungszeitpunkt bereits bezahlten Fakturen zur Berechnung von p_i zusätzlich verwendete Anzahl beträgt somit

$$\Delta m = (m - \sum_{t=0}^{-i+1} m_t) - (m - \sum_{t=0}^{-n+1} m_t) = \sum_{t=0}^{-n+1} m_t - \sum_{t=0}^{-i+1} m_t = \sum_{t=-i}^{-n+1} m_t .$$

Bei i=0 hat Δm mit $\sum_{t=0}^{-n+1} m_t$ sein Maximum und wird mit zunehmendem i kleiner, bis es bei i=n-1 den Wert m_{-n+1} erreicht. Bei der Schätzung von P_n ist kein Informationszuwachs mehr zu erzielen.

Schließlich steht noch die Bestimmung der Varianz $\hat{V}(\hat{P}_i)$ aus. Man erhält sie als

$$\hat{V}(\hat{P}_i) = \frac{p_i(1-p_i)}{m_i'-1} , \qquad (6)$$

wobei

$$m'_i = \begin{cases} m & \text{für } i = 0 \\ m - \sum_{t=0}^{-i+1} m_t & \text{für } 1 \leq i \leq -T \end{cases} \quad (7)$$

Ganz analog können auch zur Ermittlung der Betragsverteilung alle Rechnungen der Stichprobe herangezogen werden, deren Verweilzeit in der Erhebungsperiode bekannt ist. Die der Tabelle 4 entsprechende Ergebnismatrix lautet dann:

i \ t	t=0	t=-1	t=-2	t=T=-3	$\sum_{t=-i}^{-3} b_{i.t}$
i=0	$b_{0.0}$	$b_{0.,-1}$	$b_{0.,-2}$	$b_{0.,-3}$	b'_0
i=1	...	$b_{1.,-1}$	$b_{1.,-2}$	$b_{1.,-3}$	b'_1
i=2	...	...	$b_{2.,-2}$	$b_{2.,-3}$	b'_2
$r_{.t}$	$r_{.0}$	$r_{.,-1}$	$r_{.,-2}$	$r_{.,-3}$	.

Tabelle 5: Matrix der Ergebnisse einer Stichprobe zur Ermittlung der Betragsverteilung

Die Symbole haben die folgende Bedeutung:

$b_{i.t}$ ist der gesamte Betrag der in der Periode t erstellten Rechnungen, der die Verweilzeit i besitzt. Bezeichnet man mit b_{ijt} den Betrag der j-ten in t ausgestellten Rechnung, der eine Verweilzeit von i Perioden aufweist, so ist

$$b_{i.t} = \sum_{j=1}^{m_t} b_{ijt} \cdot \quad (8)$$

b'_i bezeichnet den zum Erhebungszeitpunkt t=0 beobachteten Gesamtbetrag der Stichprobe mit der Verweilzeit i. Da er sich vom tatsächlichen (nicht bekannten) Betrag mit

der Verweilzeit i unterscheidet, ist auch dieses Symbol mit einem Strich versehen. Es gilt

$$b'_i = \sum_{t=-i}^{T} b_{i.t} \cdot \tag{9}$$

$r_{.t}$ ist die Summe der Beträge aller in t erstellten Rechnungen der Stichprobe. Gibt r_{jt} den Betrag der j-ten Rechnung der Periode t an, so ist

$$r_{.t} = \sum_{j=1}^{m_t} r_{jt} \cdot \tag{10}$$

Die Betragsverteilung ergibt sich analog zur Anzahlverteilung nach

$$\hat{W}_0 = w_0 = \frac{b'_0}{\sum_{t=0}^{T} r_{.t}} \tag{11}$$

und

$$\hat{W}_i = w_i = \frac{b'_i}{\sum_{t=0}^{T} r_{.t} - \sum_{t=0}^{-i+1} r_{.t}}, \quad i=1,\ldots,n. \tag{12}$$

($\sum_{t=0}^{T} r_{.t}$ ist identisch mit dem nach $\sum_{j=1}^{m} r_j$ berechneten Betrag bei der Verhältnisschätzung im Teilabschnitt 3.5.)

Als Verhältnisschätzwerte sind $\hat{W}_0$ und $\hat{W}_i$ (i=1,...,n) leicht verzerrt. Bei genügend großem Stichprobenumfang m und großer Grundgesamtheit läßt sich die Varianz von $\hat{W}_i$ als

$$\hat{V}(\hat{W}_i) = \frac{1}{m'_i \bar{r}_i^2} \frac{\sum_{t=-i}^{T} \sum_{j=1}^{m_t} (b_{ijt} - w_i r_{jt})^2}{m'_i - 1} \tag{13}$$

schätzen, wobei

$$m'_i = \begin{cases} m & \text{für } i=0 \\ m - \sum_{t=0}^{-i+1} m_t & \text{für } 1 \leq i \leq -T \end{cases} \qquad (14)$$

und

$$\bar{r}_i = \frac{1}{m'_i} \left(\sum_{t=0}^{T} r_{.t} - \sum_{t=0}^{-i+1} r_{.t} \right) \qquad (15)$$

ist.

Die Berücksichtigung jüngerer als n Perioden alter Daten erfordert gegenüber den weiter oben vorgestellten "reinen" Schätzverfahren zusätzlich die Berechnung der Größen m_t bzw. m_t und $r_{.t}$. Der dadurch entstehende Mehraufwand erscheint jedoch tragbar; ist eine Rechnung gezogen worden, so muß ohnehin das Datum ihrer Erstellung festgehalten werden, damit die Verweilzeit berechnet werden kann. Auf der anderen Seite erspart man sich das nachträgliche Eliminieren der Fakturen, die nach t-n ausgeschrieben worden sind. Diese Einsparung dürfte die durch die Berechnung von m_t und $r_{.t}$ verursachte Mehrbelastung gewöhnlich überkompensieren.

Es darf nicht verschwiegen werden, daß durch die Modifikation der Anzahl- und Betragsschätzung eine zusätzliche Fehlerquelle auftritt, die formal aus der Art der Berechnung entsteht. Bei den reinen Verfahren ergeben sich die Schätzwerte für die Komponenten der Verweilzeitverteilungen als $p_i=a_i/m$ und $w_i=\bar{b}_i/\bar{r}$ mit einem für alle Ausprägungen von i jeweils konstanten Nenner. Dadurch wird gewährleistet, daß bei volldeckenden Spektren stets $\Sigma p_i=\Sigma w_i=1$ gilt. Bei der modifizierten Schätzung trifft das nicht mehr zu. Auch wenn tatsächlich eine volldeckende Verweilzeitverteilung vorliegt, wird sich hier sowohl Σp_i als auch Σw_i zufallsbedingt von 1 unterscheiden.

Die Abweichung wird in großen Stichproben nur geringfügig sein. Gibt die berechnete Verweilzeitverteilung jedoch Anlaß zu der Vermutung, die mit ihr durchgeführten Prognosen würden nennens-

werte systematische Vorhersagefehler aufweisen, so muß eine Korrektur angebracht werden, die Zusatzkenntnisse erfordert. Das ist z. B. der Fall, wenn $\sum_i p_i$ bzw. $\sum_i w_i$ deutlich größer als 1 ist, obwohl den Kunden nie Verzugszinsen berechnet worden sind.

Sollten in der Vergangenheit grundsätzlich alle Forderungen in voller Höhe beglichen worden sein, kann man die ermittelte Verteilung durch Multiplikation aller Elemente mit $1/\sum_i p_i$ bzw. $1/\sum_i w_i$ auf 1 normieren. Realistischer ist jedoch die Annahme, daß $\sum_i p_i$ bzw. $\sum_i w_i$ wegen Betragskürzungen und Forderungsausfällen kleiner als 1 ist. Die Verteilung ist dann auf eine entsprechend kleinere Zahl zu normieren, etwa auf

$$a^+ = 1 - \frac{\text{Anzahl der nicht liquidierten Rechnungen einer bestimmten Zeitspanne}}{\text{Anzahl der in dieser Zeitspanne insgesamt erstellten Rechnungen}}$$

bei der Anzahlverteilung und auf

$$b^+ = 1 - \frac{\text{in Anspruch genommene Skonti + Forderungsausfälle einer bestimmten Zeitspanne}}{\text{Umsatz dieser Zeitspanne}}$$

bei der Betragsverteilung, wobei man auf Erfahrungswerte zurückgreifen muß.

Durch die Normierung wird p_i zu

$$p_i^+ = p_i \frac{a^+}{\sum_{i=0}^{n} p_i}, \quad i=0,\ldots,n \tag{16}$$

und w_i zu

$$w_i^+ = w_i \frac{b^+}{\sum_{i=0}^{n} w_i}, \quad i = 0,\ldots,n\ . \tag{17}$$

Ein weiteres Problem gegenüber der Ermittlung der Verweilzeitverteilung ohne Berücksichtigung der jüngsten bezahlten Rechnungen tritt bei der Bestimmung des erforderlichen Mindeststichprobenumfangs auf, der zur Erzielung einer gewünschten Genauigkeit benötigt wird. Zieht man eine Stichprobe von m Rechnungen, so werden nur zur Schätzung von P_0 bzw. W_0 alle m Fakturen verwendet. Für die Berechnung der übrigen Komponenten der Verteilung werden mit zunehmender Verweilzeit i immer weniger Rechnungen herangezogen, wobei die jeweils benutzte Anzahl $m-\sum_{t=0}^{-i+1} m_t$ vorab nicht bekannt ist. Man kann zwar auf die gleiche Weise wie in Abschnitt 3.3 bzw. 3.6 berechnen, wieviele Rechnungen mindestens verwendet werden müssen, damit bei der Schätzung der einzelnen Komponenten P_i bzw. W_i die geforderte Genauigkeit nicht unterschritten wird; man kann aber nicht angeben, wie groß der Stichprobenumfang m sein muß, um die Einhaltung der Mindestzahl auch tatsächlich zu garantieren.

Die in 3.3 bzw. 3.6 abgeleiteten Gleichungen zur Bestimmung des Mindeststichprobenumfangs besitzen eine gewisse Sicherheitsreserve dadurch, daß sie von der ungünstigen Annahme $P_i = 0{,}5$ ausgehen, während in der Praxis meistens mit niedrigeren Werten gerechnet werden kann. Man wird deshalb in vielen Fällen zu befriedigenden Konfidenzintervallbreiten für P_i und W_i gelangen, wenn man m_{min} weiterhin auf diese Weise ermittelt. Die Genauigkeitseinbuße ist dabei besonders gering, wenn die Komponenten mit den kürzeren Verweildauern die höheren Werte aufweisen, wie es in Untersuchungen von LANGEN und seinen Mitarbeitern teilweise beobachtet wurde[1], da ja bei kleinen i fast alle m Rechnungen zur Schätzung von P_i bzw. W_i verwendet werden. Sollte es sich zeigen, daß nach Abschluß der

1 Vgl. z.B. H. Langen: Die Prognose..., a.a.O., S. 300; ders.: Betriebliche Zahlungsströme..., a.a.O., S. 262; L. Neumeyer, a.a.O., S. 65.

Ermittlung der Verweilzeitverteilung die Genauigkeit nicht ausreicht, so müssen gegebenenfalls weitere Rechnungen gezogen werden.

Bei der Zufallsentnahme müssen alle Rechnungen des ausgewählten Zeitraums die gleiche Chance besitzen, in die Stichprobe zu gelangen. Diese Bedingung ist in der Praxis recht einfach zu erfüllen, wenn die Fakturen numeriert sind und die Entnahme mit Hilfe von Zufallszahlen erfolgt. Für jede gezogene Rechnung muß dann festgestellt werden,

1. wann sie erstellt worden ist,
2. über welchen Betrag sie lautet,
3. wann sie beglichen worden ist (bzw. ob sie überhaupt schon beglichen worden ist),
4. in welcher Höhe die Bezahlung erfolgt ist.

Die Beantwortung von 1. und 2. bereitet keine Schwierigkeiten. Dagegen kann die Ermittlung der in 3. und 4. gesuchten Größen, die der Ausgangsrechnung nicht zu entnehmen sind, je nach Organisation des Rechnungswesens mehr oder weniger aufwendig sein. Im Sinne einer möglichst problemlosen Erhebung ist es günstig, wenn auf jeder Rechnung das Eingangsdatum und der liquidierte Betrag notiert werden bzw. bei EDV-geführtem Rechnungswesen diese Größen dem entsprechenden Datensatz hinzugefügt werden, sobald die Zahlung erfolgt ist. Eine solche Ergänzung erscheint zwar nicht vertretbar, wenn sie eigens zum Zwecke der Ermittlung des Liquidationsspektrums durchgeführt wird, weil dieses Vorgehen stark in die Richtung der vollständigen Erhebung der vorliegenden Fakturen zielt. Es ist aber denkbar, daß sie aus anderen Gründen, etwa im Zusammenhang mit der Zahlungsüberwachung oder aus organisatorischen Erfordernissen ohnehin praktiziert wird.

3.7.2 Ermittlung aus einer unvollständigen Stichprobe

In manchen Unternehmen werden Zweitschriften der Ausgangsrechnungen als innerbetriebliche Belege in der Reihenfolge des Geldeingangs abgelegt, nachdem Eingangsdatum und tatsächlich gezahlter Betrag auf ihnen vermerkt worden sind. Sind diese Belege fortlaufend numeriert, so bieten sie sich als Erhebungsgrundlage geradezu an. Dabei ist aber zu beachten, daß nur die bezahlten Rechnungen in die Stichprobe gelangen können. Eine Anwendung der bisher vorgestellten Schätzverfahren auf dieses Datenmaterial würde zu einer tendenziellen Überschätzung besonders der Komponenten mit kürzerer Verweilzeit führen, wie anhand der Tabelle 4 leicht einzusehen ist. Die dort mit m_t bezeichneten Werte, welche die Anzahl der in Periode t erstellten Rechnungen in der Stichprobe angeben, beziehen auch solche Rechnungen ein, die zum Erhebungszeitpunkt noch nicht beglichen worden sind und deshalb keiner der drei Größen a_{10}, a_{20} und $a_{2,-1}$ eindeutig zugeordnet werden können. Aus diesem Grund ist in der Matrix $m \geq \sum_{i=0}^{n} \sum_{t=-i}^{T} a_{it}$. Werden dagegen ausnahmslos bereits bezahlte Rechnungen entnommen, so bleiben die Zellen für a_{10}, a_{20} und $a_{2,-1}$ grundsätzlich leer, da solche Kombinationen von i und t dann nicht auftreten können. Die Stichprobe ist im Vergleich zu dem in Abschnitt 3.7.1 behandelten Fall dann unvollständig. Sie repräsentiert nicht die Verweilzeiten a l l e r Rechnungen, sondern nur die der bereits bezahlten. Die Matrix lautet jetzt

i \ t	t=0	t=-1	t=-2	t=T=-3	a_i
i=0	a_{00}	$a_{0,-1}$	$a_{0,-2}$	$a_{0,-3}$	a_0
i=1	-	$a_{1,-1}$	$a_{1,-2}$	$a_{1,-3}$	a_1
i=2	-	-	$a_{2,-2}$	$a_{2,-3}$	a_2
m_t	m_0	m_{-1}	m_{-2}	m_{-3}	m

Tabelle 6: Matrix der Ergebnisse einer Stichprobe zur Ermittlung der Anzahlverteilung, die nur aus bezahlten Rechnungen gezogen wurde

Jetzt sind a_i und m_t die tatsächlich beobachteten Randsummen in der Stichprobe und es gilt:

$$a_i = \sum_{t=0}^{T} a_{it}, \tag{1}$$

$$m_t = \sum_{i=0}^{n} a_{it}, \tag{2}$$

$$m = \sum_{i=0}^{n} a_i = \sum_{t=0}^{T} m_t = \sum_{t=0}^{T} \sum_{i=0}^{n} a_{it}. \tag{3}$$

BATSCHELET[1] und GEPPERT[2] haben für die Auswertung von Kontingenztafeln mit fehlenden Daten ein Schätzverfahren entwickelt, mit dem sich aus dieser Tabelle die Verweilzeitverteilung ermitteln läßt. Auf die Darstellung der exakten, auf einem Maximum-likelihood-Ansatz beruhenden Ableitung sei hier verzichtet, zumal man intuitiv zum gleichen Ergebnis gelangt[3]:

Die Spalte t der Matrix 6 repräsentiert die Anzahlverteilung, welche für die Umsatzforderungen der Periode t in der Stichprobe beobachtet wird. Bei zeitlicher Stabilität der Verteilung sind die Spalten bis auf Zufallsabweichungen proportional. Es liegt deshalb nahe, die Lücken in der Tabelle durch proportionale Größen auszufüllen[4]. Nach dieser Ergänzung sei a_i^+ die Anzahl der Rechnungen mit der Verweilzeit i. Offensichtlich ist

$$a_0^+ = a_0 , \tag{4}$$

---

1 Vgl. E. Batschelet: Über eine Kontingenztafel mit fehlenden Daten, Biometrische Zeitschrift 2 (1960), S. 236 ff.

2 Vgl. M. P. Geppert: Erwartungstreue plausibelste Schätzer aus dreieckig gestutzten Kontingenztafeln, Biometrische Zeitschrift 3 (1961), S. 54 ff.

3 Vgl. E. Batschelet, a.a.O., S. 239 f.

4 Ebenda.

da in der Tabelle alle Zellen der Zeile i=0 besetzt sind. Aus der Proportion

$$\frac{a_1^+}{a_0^+} = \frac{a_1}{a_0 - a_{00}} \tag{5}$$

ergibt sich wegen

$$m_0 = a_{00} \tag{6}$$

$$a_1^+ = \frac{a_1 \, a_0^+}{a_0 - m_0} \, . \tag{7}$$

Entsprechend erhält man aus

$$\frac{a_2^+}{a_0^+ + a_1^+} = \frac{a_2}{a_0 + a_1 - a_{00} - a_{0,-1} - a_{1,-1}} \tag{8}$$

wegen

$$m_0 = a_{00}$$

und

$$m_1 = a_{0,-1} + a_{1,-1} \tag{9}$$

$$a_2^+ = \frac{a_2 (a_0^+ + a_1^+)}{a_0 + a_1 - m_0 - m_{-1}}$$

$$= \frac{a_2 \sum_{k=0}^{1} a_k^+}{\sum_{k=0}^{1} (a_k - m_{-k})} \, . \tag{10}$$

a_i^+ berechnet sich demnach als

$$a_i^+ = \frac{a_i \sum_{k=0}^{i-1} a_k^+}{\sum_{k=0}^{i-1} (a_k - m_{-k})} \, , \quad i = 1, \ldots, n, \tag{11}$$

wobei man sich anhand der Tabelle 6 leicht klarmachen kann, daß

$$a_i^+ \geq a_i \ . \tag{12}$$

Der durch das proportionale Ausfüllen der Lücken entstandene vergrößerte Stichprobenumfang beträgt

$$m^+ = \sum_{i=0}^{n} a_i^+ \ , \tag{13}$$

und schließlich erhält man mit

$$\hat{p}_i = p_i = \frac{a_i^+}{m^+} \ , \ i=0,\ldots,n, \tag{14}$$

eine erwartungstreue[1] Schätzfunktion für die Anzahlverteilung der Grundgesamtheit.

Die Ermittlung der Varianz dieses Schätzers ist mit großen theoretischen Schwierigkeiten verbunden[2]. In der Praxis wird man auf sie verzichten müssen. Somit können auch keine Konfidenzintervalle für die wahren Werte der Grundgesamtheit angegeben werden.

Im Gegensatz zur Schätzung mit Berücksichtigung neuester Daten aus einer vollständigen Stichprobe wird hier bei der Berechnung jeder Komponenten der Verweilzeitverteilung durch eine für alle Komponenten gleiche Zahl, nämlich m^+, dividiert. Deshalb ist stets

$$\sum_{i=0}^{n} p_i = 1 \ . \tag{15}$$

Dieses Entnahmeverfahren läßt unbezahlte Rechnungen gar nicht erst in die Stichprobe gelangen und kann deshalb nur die Ver-

---

1 Vgl. M. P. Geppert, a.a.O., S. 66.

2 Ebenda, S. 67.

teilung der Verweilzeiten innerhalb der bezahlten Forderungen feststellen. Es ist somit noch eine Normierung auf den von den Kunden überhaupt beglichenen Prozentsatz von der Gesamtzahl der ausgefertigten Rechnungen notwendig, welcher wieder als Erfahrungswert von außen in die Schätzung eingebracht werden muß.

Die Schätzung der Betragsverteilung erfolgt analog. Bezeichnet man in der Stichprobe mit

b_i den gesamten Betrag mit der Verweilzeit i vor der Ergänzung,

b_i^+ den sich nach der proportionalen Ergänzung ergebenden Betrag mit der Verweildauer i,

$b_{..t}$ den aus den in Periode t erstellten Rechnungen tatsächlich liquidierten Betrag,

so ist $b_0^+ = b_0$ (16)

und

$$b_i^+ = \frac{b_i \sum_{k=0}^{i-1} b_k^+}{\sum_{k=0}^{i-1} (b_k - b_{..,-k})} \quad , \; i=1,\ldots,n \; . \tag{17}$$

Der Schätzwert für W_i lautet damit

$$\hat{W}_i = w_i = \frac{b_i^+}{\sum_{i=0}^{n} b_i^+} \quad , \; i=0,\ldots,n \; . \tag{18}$$

Auch hier ist wieder eine Normierung auf den von den Kunden tatsächlich gezahlten Prozentsatz des fakturierten Betrages erforderlich, der nicht aus der Stichprobe ermittelt werden kann.

Die Schätzung der Betragsverteilung nach (18) und anschließender Normierung ist plausibel. Mehr läßt sich jedoch nicht sagen. Aussagen über Erwartungswert und Varianz von $\hat{W}_i$ können nicht gemacht werden. Somit ist auch eine Intervallschätzung und die daraus abzuleitende Angabe des notwendigen Stichprobenumfangs zur Erzielung einer gewünschten Genauigkeit nicht möglich. Es ist allerdings zu vermuten, daß sich der Mindestumfang der Stichprobe in der Größenordnung desjenigen der Verhältnisschätzung bewegen wird, so daß die Abschätzung nach den dort angegebenen Gleichungen erfolgen kann.

3.8 Aufbereitung der geschätzten Verweilzeitverteilung

Ehe die ermittelte Verweilzeitverteilung zur Prognose benutzt wird, können noch zwei abschließende Manipulationen an ihr vorgenommen werden.

LANGEN schlägt vor, Verweilzeitverteilungen mit einem unregelmäßigen Erscheinungsbild vor der weiteren Verwendung zu glätten, da er die Unregelmäßigkeiten als Auswirkungen von Zufallseinflüssen ansieht[1]. Aus der Verteilung der Abbildung 7 wird auf diese Weise die geglättete Verteilung der Abbildung 8.

Die Begründung für diesen Schritt ist allerdings nicht zwingend und erfolgt vor keinem theoretischen Hintergrund, der eine bestimmte Form der Verteilung als "normal" erscheinen ließe. Liquidationsspektren ergeben sich als empirische Verteilungen und sind im Hinblick auf das wahre Zahlungsverhalten aller Kunden sicherlich mit Zufallsfehlern behaftet, ob aber Über- oder Unterschätzungen einzelner Spektralelemente vorliegen, kann ohne zusätzliche Information nicht festgestellt werden. Die Glättung ist deshalb sachlich nicht gerechtfertigt. Insbesondere wird sie von zufälligen Ausnahmen abgesehen nicht zu genaueren Prognosen führen. Praktische Untersuchungen bestätigen diese Vermutung[2].

---

1 Vgl. H. Langen: Die Prognose..., a.a.O., S. 302.

1 Vgl. R. Edin u. H. J. Schmitt, a.a.O., S. 492.

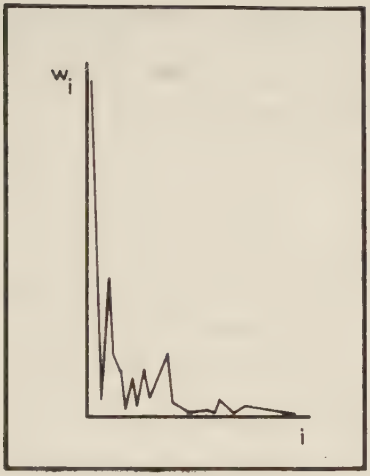

Abb. 7: Ungeglättete Verweilzeitverteilung. Nach: H. Langen: Die Prognose.., a.a.O., S. 299.

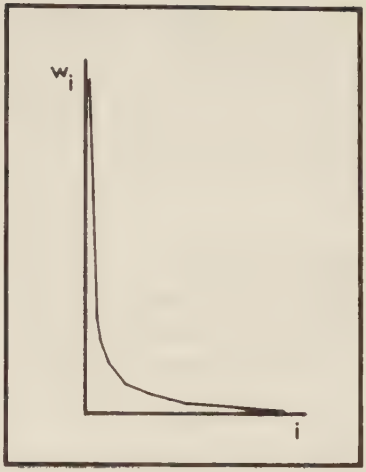

Abb. 8: Geglättete Verweilzeitverteilung. Nach: H. Langen: Die Prognose..., a.a.O., S. 304.

Die Untersuchungen haben ferner gezeigt, daß die letzten Elemente der Verweilzeitverteilung häufig recht kleine Werte aufweisen, wobei die Anzahl dieser unbedeutenden Komponenten beträchtlich sein kann. Da kleine Komponenten die Prognosewerte nur unwesentlich beeinflussen, schlagen EDIN und SCHMITT vor, die Verweilzeitverteilung zu kürzen, um den Rechenaufwand zu vermindern[1]. Diese Kürzung soll an der Anzahlverteilung erläutert werden.

Man kumuliert die Verweilzeitverteilung $(p_0\ p_1\ \dots\ p_n)$ und erhält

$$\left(\sum_{i=0}^{0} p_i \quad \sum_{i=0}^{1} p_i \quad \dots \quad \sum_{i=0}^{n} p_i\right) .$$

Die Komponente $\sum_{i=0}^{k} p_i$ der kumulierten Anzahlverteilung gibt an,

1 Vgl. R. Edin u. H. J. Schmitt, a.a.O., S. 493.

daß $100 \sum_{i=0}^{k} p_i$ Prozent der Rechnungen spätestens nach k Perioden beglichen werden. Man stellt nun fest, für welches k die Summe $\sum_{i=0}^{k} p_i$ einen bestimmten Anteil von $\sum_{i=0}^{n} p_i$, etwa 0,95, überschreitet und kappt die Verweilzeitverteilung nach diesem Element.

Prognostiziert man mit einem derart verkürzten Spektrum, so werden die Einnahmen systematisch unterschätzt. EDIN und SCHMITT empfehlen deshalb eine Normierung der verkürzten Verteilung, durch die die Summe der Komponenten auf den ursprünglichen Wert $\sum_{i=0}^{n} p_i$ zurückgebracht wird und die durch Multiplikation der Komponenten des gestutzten Spektrums mit $\sum_{i=0}^{n} p_i \; / \; \sum_{i=0}^{k} p_i$ erreicht wird.

Diese Normierung entspricht formal der des vorigen Abschnitts, unterscheidet sich jedoch von ihr inhaltlich. Dort wird die Herkunft der Einnahmeteile periodengerecht beibehalten, und es erfolgt lediglich aufgrund von Zusatzkenntnissen eine Korrektur der Höhe der Komponenten. Hier wird versucht, durch eine bewußte Überschätzung der Einnahmeteile aus jüngeren Forderungen die nichtbeachteten Teile aus älteren Forderungen zu ersetzen. Dieser Ausgleich wird nicht immer gelingen, in bestimmten Fällen, etwa bei fortlaufend wachsender Anzahl der pro Periode erstellten Rechnungen wird sogar eine zusätzliche Fehlerquelle erzeugt.

Der durch die nicht genau periodengerechte Zuweisung der Einnahmen entstehende Fehler darf allerdings gewöhnlich vernachlässigt werden[1], nicht zuletzt weil die Summe der Werte der abgeschnittenen Komponenten meistens recht klein bleibt. Das mag auch der Grund sein, weshalb dieses Problem keine weitergehende Behandlung erfahren hat. Tatsächlich ist die größenproportionale Normierung nur e i n e Möglichkeit, systematische Un-

1 Vgl. R. Edin u. H. J. Schmitt, a.a.O., S. 493.

terschätzungen zu vermeiden. Es ist nicht unbedingt erforderlich, größeren Komponenten der gestutzten Verteilung höhere absolute Beträge zuzuschlagen als kleineren. Eine denkbare Alternative wäre eine gleichmäßige Erhöhung jedes verbleibenden Elements um

$$\frac{1}{k+1} \left(\sum_{i=0}^{n} p_i - \sum_{i=0}^{k} p_i \right) .$$

Da durch das Kappen der Verweilzeitverteilung eine verursachungsgerechte Behandlung derjenigen Einnahmeteile, die später als k Perioden nach der Fakturierung eingehen, ohnehin nicht mehr möglich ist, kann man auch vollständig auf eine Zuordnung dieser Teile zu anderen Perioden verzichten und den Schätzfehler durch Addition einer konstanten Anzahl (bzw. eines konstanten Betrages bei der Betragsverteilung) ausgleichen. Diese Art der Korrektur unterstellt eine gleichbleibende Tendenz der Anzahl der Rechnungserstellungen pro Periode. Sie findet sich in Abschnitt 6.2 bei der Aufstellung eines Regressionsansatzes zur Prognose der Einnahmen aus Umsätzen wieder.

3.9 Vergleich der Verfahren

Um sich bei praktischen Untersuchungen für eines der vorgestellten Verfahren zur Ermittlung der Verweilzeitverteilung entscheiden zu können, wird sich der potentielle Anwender Klarheit darüber verschaffen wollen, wie sich die einzelnen Verfahren hinsichtlich

1. ihrer Prämissen über das Zahlungsverhalten der Kunden,
2. ihrer jeweiligen Schätzeigenschaften,
3. ihrer Benutzerfreundlichkeit

unterscheiden. Da bei allen Verfahren von denselben Annahmen über die Zahlungsweise ausgegangen wird, sind sie, nach dem Kriterium der Wirklichkeitsnähe gemessen, gleich zu beurteilen. Beträchtliche Unterschiede bestehen dagegen hinsichtlich der Schätzeigenschaften und der Praktikabilität. Sie resultie-

ren im wesentlichen aus der verschiedenartigen Bewältigung des Problems der Verklebung der DM-Stücke in den Rechnungen und den damit verbundenen Anforderungen an die Organisation des Datenmaterials. Im folgenden sei die Betrachtung auf den Vergleich der Verfahren zur Ermittlung der Betragsverteilung beschränkt, da sich die entsprechenden Aussagen auf die verhältnismäßig problemlose Schätzung der Anzahlverteilung leicht übertragen lassen.

Die NEUMEYERschen Ansätze schalten den Einfluß der Verklebung durch den direkten Zugriff auf die DM-Stücke aus. Sie erzielen dadurch erwartungstreue Schätzwerte für W_i und gestatten eine theoretisch exakte unverzerrte Schätzung der Varianz $V(\hat{W}_i)$. Diese Vorzüge werden allerdings mit einem ausgesprochen hohen Aufwand bei der Entnahme der Stichprobenelemente erkauft. Mit dem dreistufigen Ansatz wird zwar die Praktikabilität gegenüber dem einstufigen Vorgehen gesteigert, die umständliche Erstellung von Listen kumulierter Rechnungen bleibt jedoch weiterhin erforderlich; NEUMEYERs Ansicht, daß zwanzig solcher Listen ausreichen, kann aus bereits besprochenem Grunde nicht zugestimmt werden.

Die Verhältnisschätzung führt zu leicht verzerrten Schätzwerten für W_i und $V(\hat{W}_i)$ und ist NEUMEYERs Ansätzen in dieser Hinsicht unterlegen. Außerdem verlieren ihre Ergebnisse bei konstantem Stichprobenumfang mit zunehmender Varianz der Rechnungsbeträge an Genauigkeit. Dafür gestatten sie die unmittelbare Entnahme ganzer Rechnungen und umgehen die aufwendige Anfertigung von Listen mit kumulierten Beträgen. Bei Kenntnis der Streuung der Rechnungsbeträge ist ferner anders als bei der dreistufigen Ziehung zumindest eine Abschätzung des Stichprobenumfangs möglich, der eine gewünschte Konfidenzintervallbreite für W_i gewährleistet.

Erstreckt sich der Zeitraum, dem die gezogenen Rechnungen angehören können, bis zur Gegenwart, ist allerdings eine nachträgliche Reduzierung des Stichprobenumfangs notwendig, die einmal Arbeit bereitet und zum anderen in manchen Fällen die Genauigkeit der Schätzungen geringer als geplant werden läßt. Die Zusatzarbeit und den Informationsverlust kann man sich durch die Anwendung des in 3.7.1 vorgestellten Verfahrens der Ermittlung der Verweilzeitverteilung aus einer vollständigen Stichprobe ersparen. Sofern eine Normierung dabei nicht notwendig erscheint bzw. mit Hilfe externer Daten leicht möglich ist, bedeutet dieses Verfahren eine Verbesserung der "reinen" Verhältnisschätzung in den Fällen, in denen im Zeitpunkt der Erhebung nicht alle Verweilzeiten beobachtet werden können.

Die Ermittlung der Betragsverteilung aus ausschließlich bereits bezahlten Rechnungen ermöglicht höchste Arbeitsersparnis bei der Entnahme, ist jedoch in besonderem Maße auf über die Stichprobe hinausgehende Zusatzinformationen angewiesen. Unter dem Gesichtspunkt, daß die eigentliche Ziehung und nicht die sich anschließende Berechnung die Hauptarbeit der Ermittlung ausmacht, ist dieses Verfahren am günstigsten. Seine Anwendung stellt allerdings hohe Ansprüche an die Organisation des Rechnungswesens. Die theoretische Erfassung des Schätzverfahrens wird hier der Praktikabilität vollständig geopfert.

Zusammenfassend muß festgestellt werden: D a s Verfahren zur Ermittlung der Betragsverteilung gibt es nicht, da theoretische Qualität und Benutzerfreundlichkeit entgegengerichtet verlaufen. Der potentielle Anwender muß vielmehr von den organisatorischen Möglichkeiten ausgehend zu einem Kompromiß in der Erfüllung dieser Kriterien gelangen.

4 Genauigkeit der Einnahmeprognose mit Verweilzeitverteilungen

4.1 Vorbemerkungen

Bei der Vorhersage der Einnahmen aus Umsätzen nach

$$\hat{e}_t = \sum_{i=0}^{n} u_{t-i} w_i \qquad (1)$$

bzw. nach

$$\hat{e}_t = \sum_{i=0}^{n} u_{t-i} p_i \qquad (2)$$

wird eine Punktprognose abgegeben. Solch eine Angabe nur eines einzigen Prognosewertes für jede Periode gestattet zweifellos eine rechentechnisch einfache Weiterverwendung[1] von $\hat{e}_t$ für nachgelagerte Berechnungen. Daß $\hat{e}_t$ jedoch mit dem später realisierten Wert e_t tatsächlich übereinstimmt, darf man, von zufälligen Ausnahmen abgesehen, nicht erwarten. Zu viele Gründe sprechen dagegen[2].

Schon die Struktur der Prognosegleichungen (1) und (2) schließt eine exakte Übereinstimmung von $\hat{e}_t$ und e_t in der Regel aus, denn sie nimmt keine Rücksicht auf die Verklebung der DM-Stücke zu Rechnungen. Die DM-Stücke werden vielmehr behandelt, als würden sie voneinander unabhängig liquidiert[3] und der dadurch verursachte Fehler ist tendenziell umso größer, je höher der Mit-

---

1 Vgl. K. W. Rothschild: Wirtschaftsprognose - Methoden und Probleme, Berlin/Heidelberg/New York 1969, S. 166.

2 Zu den Ursachen der Prognoseabweichungen vgl. J. Schwarze: Probleme der Fehlermessung bei quantitativen ökonomischen Prognosen, ZgesStw 129 (1973), S. 539 und S. 549 f.

3 Vgl. dazu auch R. Edin: Übergangsfunktionen..., a.a.O., S. 577.

telwert und die Varianz der Rechnungsbeträge sind und je geringer die Anzahl der pro Periode erstellten Rechnungen ist[1].

Eine weitere Fehlerquelle liegt in den Parametern w_i und p_i der Prognosegleichung, die als Schätzwerte für die unbekannten Parameter der Grundgesamtheit mit Schätzfehlern behaftet sind.

In diesem Zusammenhang ist auch auf das Problem der Stabilität der Verweilzeitverteilung hinzuweisen. Die Prognosegleichungen (1) und (2) unterstellen stabile Parameter im Zeitablauf; tatsächlich kann sich das Zahlungsverhalten der Kunden in der Zukunft aber ändern und durch die vorliegenden Parameterwerte nicht mehr ausgedrückt werden.

Die erklärenden Variablen der Prognosegleichungen sind die Periodenumsätze. Soweit es sich um Vergangenheitsdaten handelt, lassen sich die u_{t-i} mit großer Genauigkeit messen. Die Gleichungen verwenden jedoch auch Umsätze zukünftiger Perioden. Diese müssen selbst prognostiziert werden und sind somit ebenfalls mit Fehlern behaftet, welche auf die Einnahmeprognose durchschlagen.

Die durch Struktur der Prognosegleichung, Schätzfehler in den Parametern und Fehler in den erklärenden Variablen verursachte Differenz zwischen Prognose und Realisation der Umsatzeinnahmen ist - in gewissem Umfang - beeinflußbar. So mildert etwa eine Vergrößerung der Basisperiode die Folgen der Nichtbeachtung der Verklebung der DM-Stücke. Die Genauigkeit der Schätzwerte für die Parameter läßt sich durch eine Vergrößerung der Stichprobe erhöhen, und bei der Prognose der zukünftigen Umsätze kann besondere Sorgfalt gezeigt werden.

Darüberhinaus existiert jedoch noch eine Fehlerquelle, die dem unmittelbaren Zugriff entzogen ist. Es handelt sich dabei um den eigentlichen Zufallsfehler, der durch die stochastische Na-

---

1 Vgl. R. Edin: Wirkungsweise..., a.a.O., S. 751 ff.

tur des Prognoseverfahrens bedingt ist und die Abweichung des realisierten Zahlungsverhaltens vom erwarteten Zahlungsverhalten bei richtig gewähltem Prognoseansatz darstellt. Er äußert sich darin, daß die tatsächlich aufgetretenen Werte von p_i bzw. w_i, mit denen sich einzelne Periodenumsätze liquidieren, von denen der Prognosegleichung, die das mittlere Zahlungsverhalten repräsentieren, unterscheiden und darf nicht mit dem Fehler in den Parametern verwechselt werden. Die Parameter sind richtig bestimmt, wenn sie das Verhalten in der G r u n d g e s a m t h e i t widerspiegeln. Abweichungen können aber auch bei Verwendung der richtigen Parameter auftreten, da die Rechnungen einer Periode nur ein Ausschnitt aus der Grundgesamtheit sind.

Das Ausmaß der Prognoseabweichung wird in einer Fehlerbetrachtung ermittelt und analysiert, welche man zweckmäßigerweise zweimal anstellt. Ex ante dient sie zur Abschätzung möglicher Abweichungen, um neben der Punktprognose auch eine Intervallprognose zu gestatten[1]. Ex post wird sie durchgeführt, um die Güte der Prognose bzw. die Richtigkeit der in die Prognosegleichung geflossenen Annahmen zu überprüfen[2]. Den unterschiedlichen Zielsetzungen entsprechend unterscheiden sich die beiden Betrachtungen. Vorab ist man auf gewisse Annahmen angewiesen, insbesondere auf solche, die das Zahlungsverhalten betreffen, und das Interesse richtet sich vorrangig auf die Abschätzung des Zufallsfehlers. Es handelt sich demnach um eine Betrachtung i n n e r h a l b des unterstellten Modells. Die bei der nachträglichen Kontrolle der Prognose beobachteten Abweichungen resultieren dagegen aus allen Fehlerquellen, enthalten also auch die Folgen falscher Annahmen. Ihre Zuordnung zu einzelnen Ursachen erfordert gewöhnlich gesonderte Untersuchungen.

Im folgenden soll zunächst untersucht werden, in welchem Umfang der Prognosefehler abschätzbar ist. Dazu wird das Zahlungsverhalten der Kunden durch ein stochastisches Modell beschrieben, welches nur ein vergleichsweise geringes Ausmaß an Annahmen und Vorkenntnissen enthält. Im Anschluß werden Maße betrachtet, die der Ex-post-Beurteilung von Prognosen dienen.

---

1 Vgl. K. W. Rothschild, a.a.O., S. 166.

2 Vgl. J. Schwarze, a.a.O., S. 538.

4.2 Ex-ante-Betrachtung des Prognosefehlers

4.2.1 Modellmäßige Beschreibung des Prognosefehlers

Der in der Periode t auftretende Prognosefehler f_t ergibt sich als Differenz zwischen den tatsächlichen Einnahmen e_t und den prognostizierten Einnahmen $\hat{e}_t$ [1]:

$$f_t = e_t - \hat{e}_t = e_t - \sum_{i=0}^{n} u_{t-i} w_i \quad . \tag{1}$$

e_t ist zum Zeitpunkt der Prognoseabgabe noch nicht bekannt und wird im Rahmen des stochastischen Modells als Realisation einer Zufallsvariablen angesehen, die das Symbol E_t erhalten soll. Zur Berechnung von $\hat{e}_t$ werden die Stichprobenwerte w_i (i=0,...,n) verwendet, die Ausprägungen der Zufallsvariablen $\hat{W}_i$ (i=0,...,n) darstellen. Der Prognosefehler ist damit ebenfalls eine Zufallsgröße:

$$F_t = E_t - \sum_{i=0}^{n} u_{t-i} \hat{W}_i \quad . \tag{2}$$

Er läßt sich gemäß

$$F_t = (E_t - \sum_{i=0}^{n} u_{t-i} W_i) + (\sum_{i=0}^{n} u_{t-i} W_i - \sum_{i=0}^{n} u_{t-i} \hat{W}_i) \tag{3}$$

in zwei Komponenten aufspalten[2]:

- Die Fehlerkomponente $(E_t - \Sigma u_{t-i} W_i)$ entsteht dadurch, daß das Zahlungsverhalten der Kunden in der Periode t vom mittleren Zahlungsverhalten abweichen wird. Es handelt sich hier um denjenigen Zufallsfehler, welcher der stochastischen Natur des Modells entspringt und nicht verringert werden kann.

---

1 Vgl. R. Edin: Wirkungsweise..., a.a.O., S. 750.

2 Vgl. J. Kmenta: Elements of Econometrics, New York 1971, S. 139 f.

- $(\Sigma u_{t-i}W_i - \Sigma u_{t-i}\hat{W}_i)$ gibt den Teil des Prognosefehlers an, der aufgrund der Verwendung von Schätzwerten $\hat{W}_i = w_i (i=0,\ldots,n)$ anstelle der unbekannten Parameter W_i auftritt. Das Ausmaß dieser Abweichung hängt maßgeblich vom Umfang der zur Ermittlung der w_i gezogenen Stichprobe ab.

Der Erwartungswert von F_t lautet

$$E(F_t) = E(E_t - \sum_{i=0}^{n} u_{t-i}W_i) + E(\sum_{i=0}^{n} u_{t-i}W_i - \sum_{i=0}^{n} u_{t-i}\hat{W}_i) \quad (4)$$

und die Varianz

$$V(F_t) = V(E_t - \sum_{i=0}^{n} u_{t-i}W_i) + V(\sum_{i=0}^{n} u_{t-i}W_i - \sum_{i=0}^{n} u_{t-i}\hat{W}_i)$$

$$= V(E_t) + V(\sum_{i=0}^{n} u_{t-i}\hat{W}_i) \; . \quad (5)$$

(Die Kovarianz $C(E_t, \Sigma u_{t-i}\hat{W}_i)$ verschwindet, weil die beiden Fehlerkomponenten als Ergebnisse zweier voneinander stochastisch unabhängiger Zufallsexperimente aufgefaßt werden können.)

Da sich die Additivität der beiden Fehlerkomponenten auch auf den Erwartungswert und die Varianz erstreckt, bietet sich für den weiteren Verlauf der Analyse eine getrennte Betrachtung an.

Zunächst sei die Variable E_t untersucht. Sie wird vom tatsächlichen Zahlungsverhalten in der Periode t bestimmt. Die Zahlungsweise der Kunden wiederum wird durch eine derartige Vielzahl von Größen[1] beeinflußt, daß der Versuch einer vollständigen Beschreibung erfolglos bleiben muß. Man ist deshalb gezwungen, die Wirklichkeit durch ein Modell vereinfachend zu beschreiben. Dazu werden die folgenden Annahmen getroffen:

---

1 Vgl. H. Kossbiel, a.a.O., S. 66 ff.

1. Die in einer Periode erstellten Rechnungen können als Zufallsstichprobe aus der alle Rechnungen umfassenden Grundgesamtheit interpretiert werden. Es wird unterstellt, daß sich der Mittelwert und die Varianz der Beträge der Rechnungen einer Periode und die tatsächliche Verweilzeitverteilung, nach der diese Beträge liquidiert werden, von den entsprechenden Werten der Grundgesamtheit nur in der Größenordnung von zufälligen Abweichungen unterscheiden. Ferner soll die Bezahlung der einzelnen Rechnungen voneinander unabhängig erfolgen. Die 1. Annahme schließt die Forderung der Stationarität der Verweilzeitverteilung ein.

2. Jede Rechnung wird in einem Zahlungsakt oder gar nicht beglichen, Teilzahlung tritt also nicht auf.

3. Die Anzahl der pro Periode erstellten Rechnungen ist groß, jedoch sehr klein gemessen an der Grundgesamtheit.

4. Die Prognose erfolgt mit der Betragsverteilung, die nach der Verhältnisschätzung ermittelt wurde.

Da die Annahmen 2 - 4 eher technische Voraussetzungen darstellen, werden sich Bedenken wohl vorwiegend gegen die 1. Prämisse erheben. KOSSBIEL hat in einer umfangreichen theoretischen Analyse des Zahlungsverhaltens Möglichkeiten aufgezeigt, welche diese Annahme in vielen Fällen rechtfertigen[1]. Er räumt jedoch ein, daß es sich dabei um eine reine abstrahierende Deduktion handelt, die nur in einer empirischen Untersuchung bestätigt werden kann[2]. Allgemein läßt sich feststellen, daß die Annahme tendenziell umso besser erfüllt ist, je größer die Anzahl der pro

---

1 Vgl. H. Kossbiel, a.a.O., S. 80 ff.
2 Ebenda, S. 14.

Periode anfallenden Rechnungen ist[1] und je weniger die Struktur des Kundenstammes sich im Zeitablauf verändert[2].

Der Modellansatz soll die praktische Berechnung eines Konfidenzintervalls für die Einnahmeprognose ermöglichen. Er muß deshalb auf die explizite Berücksichtigung der Rechnungsbeträge r_j der einzelnen Fakturen, aus denen sich die Periodenumsätze zusammensetzen, verzichten. Es wird sich zeigen, daß diese Forderung zu Genauigkeitseinbußen bei der Intervallprognose führt. Sie läßt sich jedoch nicht vermeiden, weil die Einzelbeträge der Rechnungen zukünftiger Perioden gewöhnlich noch nicht bekannt sind. Bei großem Geschäftsvolumen kann eine Finzelberücksichtigung aus Kostengründen auch gar nicht erwünscht sein. Die Ermittlung des Konfidenzintervalls soll vielmehr ohne großen Aufwand aus Daten erfolgen, welche unter der Gültigkeit der Prämissen im Zeitpunkt der Prognoseabgabe bereits vorliegen. Dazu wird die folgende Überlegung angestellt:

Bezeichnet man mit $b_{ij,t-i}$ denjenigen Betrag aus der j-ten Rechnung der Periode t-i, der die Verweilzeit i besitzt, und mit m_{t-i} die Anzahl der in t-i insgesamt erstellten Rechnungen, so läßt sich e_t auch als

$$e_t = \sum_{i=0}^{n} \sum_{j=1}^{m_{t-i}} b_{ij,t-i} \qquad (6)$$

ausdrücken. Die Beträge $b_{ij,t-i}$ $(j=1,\ldots,m_{t-i})$ der Periode t-i sollen als Zufallsstichprobe vom Umfang m_{t-i} aus der Grundgesamtheit aller b_{ij} $(j=1,\ldots,M)$ aufgefaßt werden. Man kann sich dann jedes $b_{ij,t-i}$ als Ergebnis einer Ziehung aus der Grundgesamtheit vorstellen. Bei Gültigkeit der beiden ersten Annahmen läßt sich $b_{ij,t-i}$ als Ausprägung einer Zufallsvariablen B_i mit dem Erwartungswert

---

1 Vgl. H. Langen: Die betriebswirtschaftliche Dispositionsrechnung im Dienste der Unternehmensführung, in: R. W. Stöhr (Hrsg.): Unternehmensführung auf neuen Wegen, Wiesbaden 1967, S. 242.

2 Vgl. H. Kossbiel, a.a.O., S. 86 ff.

$$E(B_i) = \bar{B}_i = \frac{1}{M} \sum_{j=1}^{M} b_{ij} \quad , \tag{7}$$

der Varianz

$$V(B_i) = \frac{1}{M} \sum_{j=1}^{M} (b_{ij} - \bar{B}_i)^2 \tag{8}$$

und einer unbekannten Verteilungsform ansehen. Gleichwertig kann man $b_{ij,t-i}$ auch als Realisation einer Zufallsvariablen $B_{ij,t-i}$ auffassen, die die gleiche Verteilung wie B_i besitzt, für die also insbesondere auch

$$E(B_{ij,t-i}) = E(B_i) = \bar{B}_i \tag{7a}$$

und

$$V(B_{ij,t-i}) = V(B_i) \tag{8a}$$

gilt. Die Einnahmen in t lassen sich dann als

$$E_t = \sum_{i=0}^{n} \sum_{j=1}^{m_{t-i}} B_{ij,t-i} \tag{9}$$

beschreiben. Ihr Erwartungswert beträgt

$$E(E_t) = \sum_{i=0}^{n} \sum_{j=1}^{m_{t-i}} E(B_{ij,t-i}) = \sum_{i=0}^{n} \sum_{j=1}^{m_{t-i}} \bar{B}_i = \sum_{i=0}^{n} m_{t-i} \bar{B}_i \tag{10}$$

oder, wegen

$$W_i = \frac{\bar{B}_i}{\bar{R}} \quad ,$$

$$E(E_t) = \sum_{i=0}^{n} m_{t-i} \bar{R} W_i \quad . \tag{11}$$

Der Erwartungswert der Fehlerkomponente $(E_t - \sum_{i=0}^{n} u_{t-i} W_i)$

ergibt sich somit als

$$E(E_t - \sum_{i=0}^{n} u_{t-i} W_i) = \sum_{i=0}^{n} m_{t-i} \bar{R} W_i - \sum_{i=0}^{n} u_{t-i} W_i \quad . \tag{12}$$

Bezeichnet man mit $r_{j,t-i}$ den Betrag der j-ten in Periode t-i erstellten Rechnung und mit $\bar{r}_{t-i}$ den Mittelwert der $r_{j,t-i}$ $(j=1,\ldots,m_{t-i})$, d. h.

$$\bar{r}_{t-i} = \frac{1}{m_{t-i}} \sum_{j=1}^{m_{t-i}} r_{j,t-i} \,, \tag{13}$$

und bedenkt man, daß

$$u_{t-i} = m_{t-i}\,\bar{r}_{t-i} \,, \tag{14}$$

so läßt sich (12) umformen zu

$$E(E_t - \sum_{i=0}^{n} u_{t-i}W_i) = \sum_{i=0}^{n} m_{t-i}\bar{R}W_i - \sum_{i=0}^{n} m_{t-i}\bar{r}_{t-i}W_i$$

$$= \sum_{i=0}^{n} m_{t-i}W_i(\bar{R} - \bar{r}_{t-i}) \,. \tag{15}$$

$(\bar{R}-\bar{r}_{t-i})$ ist die Abweichung des Stichprobenmittels vom Mittelwert der Grundgesamtheit. Je größer der Stichprobenumfang m_{t-i} ist, je mehr Rechnungen also in t-i ausgestellt werden, umso geringer fällt diese Abweichung tendenziell aus. Bei großem m_{t-i} ist $(\bar{R}-\bar{r}_{t-i})$ annähernd Null[1]:

$$(\bar{R} - \bar{r}_{t-i}) \triangleq 0 \,. \tag{16}$$

Damit ist auch

$$\sum_{i=0}^{n} m_{t-i}W_i(\bar{R}-\bar{r}_{t-i}) \triangleq 0 \,, \tag{17}$$

d. h.: Bei der Prognose mit Hilfe der Betragsverteilung der Grundgesamtheit tritt eine Verzerrung von $\Sigma m_{t-i}W_i(\bar{R}-\bar{r}_{t-i})$ auf, die bei großer Anzahl von Rechnungserstellungen pro Periode vernachlässigt werden kann.

---

1 Vgl. W. G. Cochran, a.a.O., S. 47.

Die Varianz der Fehlerkomponente $(E_t - \Sigma u_{t-i} W_i)$ schließlich lautet

$$V(E_t - \sum_{i=0}^{n} u_{t-i} W_i) = V(E_t)$$

$$= \sum_{i=0}^{n} \sum_{j=1}^{m_{t-i}} V(B_i)$$

$$= \sum_{i=0}^{n} m_{t-i} V(B_i) \quad . \tag{18}$$

Die Gestalt der Verteilung von E_t und damit auch die von $(E_t - \Sigma u_{t-i} W_i)$ wird durch die Verteilungen der B_i (i=0,...,n) bestimmt. Wie man sich die Verteilungsform der Zufallsvariablen B_i in der Grundgesamtheit vorstellen kann, soll an einem einfachen hypothetischen Beispiel gezeigt werden. Aus didaktischen Gründen sei zunächst Unabhängigkeit von Verweilzeit und Betrag und vollständige Bezahlung aller Rechnungen unterstellt, d. h.

$$W_i = P_i \text{ für alle } i, \tag{19}$$

$$\sum_{i=0}^{n} W_i = \sum_{i=0}^{n} P_i = 1 \text{ und} \tag{20}$$

$$B_i = RX_i, \tag{21}$$

wobei X_i wieder die Null-Eins-Variable mit $E(X_i)=P_i$ und R die Zufallsvariable "Rechnungsbetrag" ist. Der Übersichtlichkeit wegen seien ferner die Rechnungen nach ihrer Höhe in Betragsklassen eingeteilt.

In der Grundgesamtheit gelte folgende relative Häufigkeitsverteilung:

Betragsklasse k	Rechnungsbetrag R	relative Häufigkeit der Rechnungen dieser Klasse $f_{k,R}$
1	bis 1 000 DM	0,20
2	über 1 000 bis 2 000 DM	0,40
3	über 2 000 bis 3 000 DM	0,15
4	über 3 000 bis 4 000 DM	0,07
5	über 4 000 bis 5 000 DM	0,05
6	über 5 000 bis 6 000 DM	0,05
7	über 6 000 bis 7 000 DM	0,04
8	über 7 000 DM	0,04

Tabelle 7: Beispiel für die Verteilung von Rechnungsbeträgen in der Grundgesamtheit

Abbildung 9 soll diese Verteilung verdeutlichen:

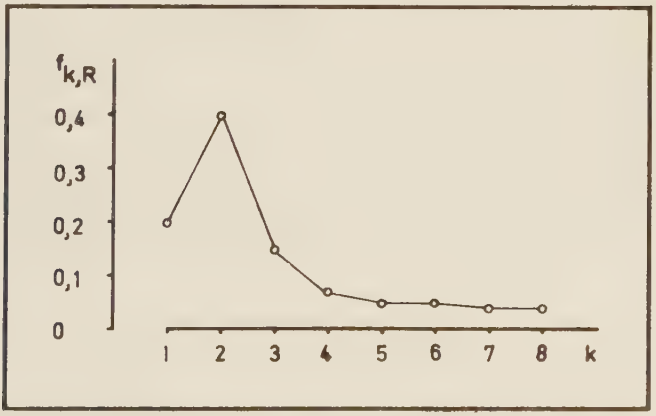

Abb. 9: Graphische Darstellung der relativen Häufigkeiten des Beispiels[1]

1 Es gelten nur die herausgehobenen Punkte.

Wird die j-te Rechnung nach i Perioden bezahlt, so ist $b_{ij}=r_j$ und fällt in eine der angegebenen Betragsklassen. Die Wahrscheinlichkeit dafür beträgt unter den obigen Annahmen P_i, und die Wahrscheinlichkeit dafür, daß B_i gerade in die k-te Klasse fällt, ist $f_{k,R}P_i$. Darüber hinaus kann B_i auch den Wert $b_{ij}=0$ annehmen, der eine zusätzliche Betragsklasse bildet, die mit 0 bezeichnet werden soll. Da $b_{ij}=0$ mit der Wahrscheinlichkeit $1-P_i$ auftritt, ergibt sich für B_i die in Tabelle 8 wiedergegebene relative Häufigkeitsverteilung.

Betragsklasse k	bezahlter Betrag B_i	relative Häufigkeit der Rechnungen dieser Klasse $f_{k,B}$
0	genau 0 DM	$1-P_i$
1	über 0 bis 1 000 DM	$0,20\ P_i$
2	über 1 000 bis 2 000 DM	$0,40\ P_i$
3	über 2 000 bis 3 000 DM	$0,15\ P_i$
4	über 3 000 bis 4 000 DM	$0,07\ P_i$
5	über 4 000 bis 5 000 DM	$0,05\ P_i$
6	über 5 000 bis 6 000 DM	$0,05\ P_i$
7	über 6 000 bis 7 000 DM	$0,04\ P_i$
8	über 7 000 DM	$0,04\ P_i$

Tabelle 8: Beispiel für die Verteilung von bezahlten Einzelbeträgen in der Grundgesamtheit

Die Abbildung 10 zeigt, in welcher Weise die Verteilung der Variablen B_i von P_i abhängt. Bei $P_i=1$ ist sie unter den gesetzten Annahmen mit der Verteilung von R identisch. Je kleiner P_i wird, umso schiefer wird sie. Anhand der Abbildung 10 kann man sich leicht klarmachen, daß diese Aussage für beliebige Formen der Verteilung von R gilt!

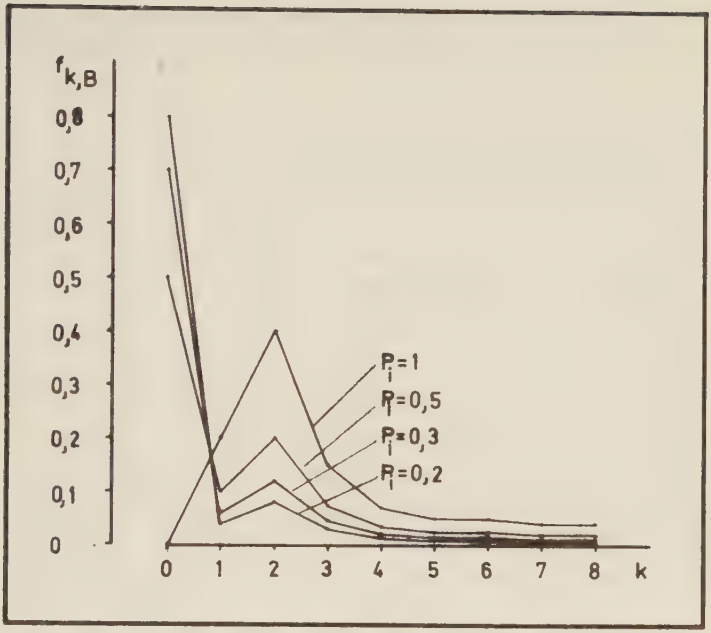

Abb. 10: Verteilung von B_i für verschiedene P_i [1]

Ist die Anzahl der pro Periode erstellten Rechnungen genügend groß, so ist $\sum_i \sum_j B_{ij,t-i}$ annähernd normalverteilt. Die Annäherung erfolgt dabei in zwei Schritten:

$\sum_{j=1}^{m_{t-i}} B_{ij,t-i}$ stellt das Ergebnis von m_{t-i} Ziehungen aus der Grundgesamtheit dar und ist bei großem m_{t-i} bereits selbst annähernd normalverteilt[2]. Damit schon auf dieser Stufe eine gute Approximation erzielt wird, sind allerdings - besonders bei kleinen P_i - wegen der starken Schiefe der Verteilungen der B_i extrem große

1 Der besseren Übersicht wegen wurde auf eine stärkere Markierung der diskreten Häufigkeiten verzichtet.

2 Vgl. E. Kreyszig, a.a.O., S. 153.

Werte von m_{t-i} erforderlich. Nach einer Faustregel sollte m_{t-i} für eine ausreichende Approximation größer als $25G_i^2$ sein[1], wobei

$$G_i = \frac{\frac{1}{M}\sum_{j=1}^{M}(b_{ij}-\bar{B}_i)^3}{\left[\frac{1}{M}\sum_{j=1}^{M}(b_{ij}-\bar{B}_i)^2\right]^{\frac{3}{2}}}$$

das FISHERsche Schiefemaß, ist, welches bei der Ermittlung der Verweilzeitverteilung mitberechnet werden kann.

Durch die Addition der Summen $\sum_{j=1}^{m_{t-i}} B_{ij,t-i}$ zu $\sum_{i=0}^{n}\sum_{j=1}^{m_{t-i}} B_{ij,t-i}$ erfolgt nach dem zentralen Grenzwertsatz von LJAPUNOW eine weitere Annäherung an die Normalverteilung[2]. Sie fällt umso besser aus, je größer n, die Länge der Verweilzeitverteilung, ist und je gleichmäßiger die Beiträge der $\sum_j B_{ij,t-i}$ zur Gesamtsumme $\sum_i\sum_j B_{ij,t-i}$ sind, je weniger sich also die Komponenten P_i unterscheiden.

Da lange Verweilzeitverteilungen tendenziell mit kleinen Werten von P_i einhergehen, findet ein gewisser Ausgleich statt: Je schlechter die Approximation auf der 1. Stufe ist, umso besser ist sie (tendenziell) auf der 2. Stufe. Allgemein läßt sich somit etwas grob sagen, daß die Approximation der Verteilung von E_t durch die Normalverteilung umso besser ausfällt, je größer $\sum_{i=0}^{n} m_{t-i}$ ist.

Hebt man schließlich die Annahmen "Unabhängigkeit von Betrag und Verweilzeit" und "volldeckendes Spektrum" auf, so lassen sich die relativen Häufigkeiten der Klassen 1 bis 8 nicht mehr

---

1 Vgl. W. G. Cochran, a.a.O., S. 60.

2 Vgl. etwa L. Schmetterer: Einführung in die mathematische Statistik, Zweite, verbesserte und wesentlich erweiterte Auflage, Wien/New York 1966, S. 126 ff; S. S. Wilks: Mathematical Statistics, Second printing with corrections, New York/London 1963, S. 257 ff.

als $f_{k,B} = f_{k,R}\, P_i$ ermitteln. Die Verteilungen der B_i sind dann in diesen Klassen für verschiedene Werte von P_i nicht mehr ähnlich; an den grundsätzlichen Gedanken zur Approximation ändert sich jedoch nichts.

Man kann somit feststellen: Bei großer Anzahl von Rechnungserstellungen pro Periode folgt die Fehlerkomponente $(E_t - \Sigma u_{t-i} W_i)$ in Annäherung einer Normalverteilung mit dem Erwartungswert Null und der Varianz $\Sigma m_{t-i} V(B_i)$.

Nun sei die Komponente $\Sigma u_{t-i}(W_i - \hat{W}_i)$ betrachtet. Im Zusammenhang mit der Ermittlung der Verweilzeitverteilung wurde bereits angegeben, daß $\hat{W}_i$ bei großem Stichprobenumfang m ein annähernd normalverteilter und praktisch unverzerrter Schätzwert für W_i ist,

$$E(\hat{W}_i) \doteq W_i \ , \tag{22}$$

dessen Varianz sich näherungsweise durch

$$V(\hat{W}_i) \doteq \frac{1}{m\bar{R}^2} \frac{1}{M} \sum_{j=1}^{M} (b_{ij} - W_i r_j)^2 \tag{23}$$

oder gleichwertig durch

$$V(\hat{W}_i) \doteq \frac{1}{m\bar{R}^2} \left[V(B_i) - 2\, W_i C(B_i, R) + W_i^2 V(R) \right] \tag{24}$$

beschreiben läßt. Damit ist auch die Komponente $\Sigma u_{t-i}(W_i - \hat{W}_i)$ bei großem m approximativ normalverteilt mit dem Erwartungswert

$$E\left[\sum_{i=0}^{n} u_{t-i}(W_i - \hat{W}_i) \right] \doteq 0 \tag{25}$$

und der Varianz

$$V\left[\sum_{i=0}^{n} u_{t-i}(W_i - \hat{W}_i) \right] \doteq \sum_{i=0}^{n} u_{t-i}^2 V(\hat{W}_i) + \sum_{i=0}^{n} \sum_{\substack{k=0 \\ k \neq i}}^{n} u_{t-i} u_{t-k} C(\hat{W}_i, \hat{W}_k). \tag{26}$$

Die Kovarianz $C(\hat{W}_i, \hat{W}_k)$ ergibt sich dabei in Analogie zu (23) als

$$C(\hat{W}_i, \hat{W}_k) \doteq \frac{1}{m\bar{R}^2} \frac{1}{M} \sum_{j=1}^{M} (b_{ij} - W_i r_j)(b_{kj} - W_k r_j). \tag{27}$$

Zu einer der geläufigeren Schreibweise (24) analogen Form der Kovarianz gelangt man durch einige Umformungen von (27):

$$C(\hat{W}_i,\hat{W}_k) \doteq \frac{1}{m\bar{R}^2}\frac{1}{M}\sum_{j=1}^{M}(b_{ij}b_{kj} - W_i b_{kj}r_j - W_k b_{ij}r_j + W_i W_k r_j^2)$$

$$\doteq \frac{1}{m\bar{R}^2}\left[\frac{1}{M}\sum_{j=1}^{M} b_{ij}b_{kj} - \frac{W_i}{M}\sum_{j=1}^{M} b_{kj}r_j - \frac{W_k}{M}\sum_{j=1}^{M} b_{ij}r_j + \frac{W_i W_k}{M}\sum_{j=1}^{M} r_j^2\right] \qquad (28)$$

oder, unter Berücksichtigung der Identität

$$-\bar{B}_i\bar{B}_k + W_i\bar{B}_k\bar{R} + W_k\bar{B}_i\bar{R} - W_i W_k\bar{R}^2 = 0,$$

$$C(\hat{W}_i,\hat{W}_k) \doteq \frac{1}{m\bar{R}^2}\left[\left(\frac{1}{M}\sum_{j=1}^{M} b_{ij}b_{kj} - \bar{B}_i\bar{B}_k\right) - W_i\left(\frac{1}{M}\sum_{j=1}^{M} b_{kj}r_j - \bar{B}_k\bar{R}\right)\right.$$

$$\left. - W_k\left(\frac{1}{M}\sum_{j=1}^{M} b_{ij}r_j - \bar{B}_i\bar{R}\right) + W_i W_k\left(\frac{1}{M}\sum_{j=1}^{M} r_j^2 - \bar{R}^2\right)\right]$$

$$\doteq \frac{1}{m\bar{R}^2}\left[C(B_i,B_k) - W_i C(B_k,R) - W_k C(B_i,R) + W_i W_k V(R)\right] . \qquad (29)$$

Wenn keine Teilzahlung auftritt, gilt für jede Rechnung j

$$b_{ij}b_{kj} = 0 \quad \text{für } i \neq k$$

und damit auch

$$\sum_{j=1}^{M} b_{ij}b_{kj} = 0 \quad \text{für } i \neq k . \qquad (30)$$

Dann vereinfacht sich $C(B_i,B_k)$ zu

$$C(B_i,B_k) = -\bar{B}_i\bar{B}_k$$

und $C(\hat{W}_i,\hat{W}_k)$ zu

$$C(\hat{W}_i,\hat{W}_k) \doteq \frac{1}{m\bar{R}^2}\left[W_iW_kV(R) - B_iB_k - W_iC(B_k,R) - W_kC(B_i,R)\right] . \qquad (31)$$

$C(\hat{W}_i,\hat{W}_k)$ wird also ausschließlich durch Größen bestimmt, für die bei der Ermittlung der Betragsverteilung mittels der Verhältnisschätzung bereits Schätzwerte vorliegen, und läßt sich somit ebenfalls schätzen. Für eine praktische Berechnung empfiehlt sich ausgehend von (28) unter Berücksichtigung von (30) die Form

$$\hat{C}(\hat{W}_i,\hat{W}_k) = \frac{1}{m\bar{r}^2(m-1)}\left(w_iw_k\sum_{j=1}^{m} r_j^2 - w_i\sum_{j=1}^{m} b_{kj}r_j - w_k\sum_{j=1}^{m} b_{ij}r_j\right). \qquad (32)$$

Die Ergebnisse der Untersuchung der Fehlerkomponenten seien abschließend kurz zusammengefaßt:

- Bei großer Anzahl von Rechnungserstellungen pro Periode und der Ermittlung der Betragsverteilung aus einer großen Stichprobe sind beide Komponenten und damit auch der gesamte Prognosefehler F_t annähernd normalverteilt.

- Unter diesen und den übrigen getroffenen Annahmen beträgt der Erwartungswert von F_t

$$E(F_t) = E(E_t - \sum_{i=0}^{n} u_{t-i}W_i) + E(\sum_{i=0}^{n} u_{t-i}(W_i - \hat{W}_i)) \doteq 0 .$$

  $\hat{e}_t$ ist dann eine praktisch erwartungstreue Prognose für die Einnahmen aus Umsätzen in t.

- Die Varianz des Prognosefehlers läßt sich bei Gültigkeit der Prämissen als

$$V(F_t) = V(E_t) + V\left(\sum_{i=0}^{n} u_{t-i}\hat{W}_i\right)$$

$$\doteq \sum_{i=0}^{n} m_{t-i}V(B_i) + \sum_{i=0}^{n} u_{t-i}^2V(\hat{W}_i) + \sum_{i=0}^{n}\sum_{\substack{k=0\\k\neq i}}^{n} u_{t-i}u_{t-k}C(\hat{W}_i,\hat{W}_k)$$

angeben und kann aus der zur Ermittlung der Betragsverteilung gezogenen Stichprobe geschätzt werden.

4.2.2 Einflußgrößen der Prognosegenauigkeit

Welche Größen auf $V(F_t)$ und damit auf die Prognosegenauigkeit einwirken, läßt sich am besten an dem Spezialfall "Unabhängigkeit von Rechnungsbetrag und Verweilzeit" aufzeigen. Im Zusammenhang mit der Bestimmung des Mindeststichprobenumfangs wurde gezeigt, daß sich B_i dann als

$$B_i = (1 - c_i)\, RX_i \qquad (1)$$

ausdrücken läßt, wobei

c_i der Anteil der Einnahmeminderung (Skonto u.ä.) bei Bezahlung nach i Perioden,

R die Zufallsvariable "Rechnungsbetrag" mit $E(R) = \bar{R}$ und

X_i eine Null-Eins-Variable mit $E(X_i) = P_i$ und $V(X_i) = P_i(1-P_i)$ ist.

Ferner sind dann R und X_i stochastisch unabhängig und es gilt $W_i = (1-c_i)P_i$.

Die Varianz von B_i lautet in diesem Fall[1]

$$\begin{aligned} V(B_i) &= V\left[(1-c_i)RX_i\right] \\ &= (1-c_i)^2 V(RX_i) \\ &= (1-c_i)^2 \left[\left[E(X_i)\right]^2 V(R) + \left[E(R)\right]^2 V(X_i) + V(R)V(X_i)\right] \\ &= (1-c_i)^2 \left[P_i^2 V(R) + \bar{R}^2 P_i(1-P_i) + V(R)P_i(1-P_i)\right] \\ &= (1-c_i)^2 \left[\bar{R}^2 P_i(1-P_i) + V(R)P_i\right]. \end{aligned} \qquad (2)$$

1 Zur Varianz des Produkts zweier unabhängiger Variablen vgl. etwa A. M. Mood, F. A. Graybill u. D. C. Boes: Introduction to the Theory of Statistics, Third edition, Tokyo 1974, S. 180.

Die Varianz der Fehlerkomponente $(E_t - \Sigma u_{t-i} W_i)$ beträgt somit[1]

$$V(E_t) = \sum_{i=0}^{n} m_{t-i} (1-c_i)^2 \left[\bar{R}^2 P_i (1-P_i) + V(R) P_i\right] . \qquad (3)$$

Nun zur Fehlerkomponente $\Sigma u_{t-i}(W_i - \hat{W}_i)$: Bei Unabhängigkeit von Rechnungsbetrag und Verweilzeit ergab sich[2]

$$V(\hat{W}_i) \doteq \frac{1}{m} P_i (1-P_i) (1-c_i)^2 \left(1 + \frac{V(R)}{\bar{R}^2}\right) . \qquad (4)$$

Entsprechend erhält man in diesem Fall die Kovarianz $C(\hat{W}_i, \hat{W}_k)$ ausgehend von

$$C(\hat{W}_i, \hat{W}_k) \doteq \frac{1}{m\bar{R}^2} \frac{1}{M} \sum_{j=1}^{M} (b_{ij} - W_i r_j)(b_{kj} - W_k r_j) \qquad (5)$$

als

$$\begin{aligned} C(\hat{W}_i, \hat{W}_k) &\doteq \frac{1}{m\bar{R}^2} \frac{1}{M} \sum_{j=1}^{M} r_j^2 \left(\frac{b_{ij}}{r_j} - W_i\right)\left(\frac{b_{kj}}{r_j} - W_k\right) \\ &\doteq \frac{1}{m\bar{R}^2} \frac{1}{M} \sum_{j=1}^{M} r_j^2 \left[(1-c_i) x_{ij} - W_i\right]\left[(1-c_k) x_{kj} - W_k\right] \\ &\doteq \frac{1}{m\bar{R}^2} E\Big\{R^2 \Big[(1-c_i)(1-c_k) X_i X_k - (1-c_i) W_k X_i \\ &\qquad - (1-c_k) W_i X_k + W_i W_k\Big]\Big\} \\ &\doteq \frac{1}{m\bar{R}^2} \Big\{E(R^2) \Big[(1-c_i)(1-c_k) E(X_i X_k) - (1-c_i) W_k E(X_i) \\ &\qquad - (1-c_k) W_i E(X_k) + W_i W_k\Big]\Big\} . \qquad (6) \end{aligned}$$

---

1 Vgl. dazu auch H. Benishay: A Stochastic Model of Credit Sales Debt, Journal of the American Statistical Association 61 (1966), S. 1010 ff; ders.: Parameters and Relations of Stochastically Lagged and Disaggregative Time Series, Econometrica 36 (1968), S. 155 ff. BENISHAY hat auf etwas andere Weise ein sehr allgemeines, auch für Lagerhaltungsvorgänge anwendbares stochastisches Modell mit zufallsgesteuerter Anzahl von Inputereignissen (hier: m_{t-i}) aufgestellt, das unter den hier getroffenen Annahmen zur gleichen Fehlervarianz gelangt.

2 Vgl. S. 56.

Darin gilt

$$E(X_i) = P_i$$

$$E(X_k) = P_k$$

und, wenn keine Teilzahlung vorliegt,

$$E(X_i X_k) = 0,$$

da X_i und X_k dann nicht gleichzeitig von Null verschieden sein können. Somit ist

$$C(\hat{W}_i, \hat{W}_k) \doteq \frac{1}{m\bar{R}^2} \left\{ E(R^2) \left[W_i W_k - (1-c_i) P_i W_k - (1-c_k) W_i P_k \right] \right\}$$

$$\doteq \frac{1}{m\bar{R}^2} \left\{ E(R^2) \left[(1-c_i)(1-c_k) P_i P_k - (1-c_i)(1-c_k) P_i P_k \right. \right.$$

$$\left. \left. - (1-c_i)(1-c_k) P_i P_k \right] \right\}$$

$$\doteq - \frac{1}{m} P_i P_k (1-c_i)(1-c_k) \left[1 + \frac{V(R)}{\bar{R}^2} \right] . \qquad (7)$$

Die Varianz der Fehlerkomponente $\Sigma u_{t-i} (W_i - \hat{W}_i)$ beträgt für diesen Sonderfall also

$$V\left(\sum_{i=0}^{n} u_{t-i} \hat{W}_i \right) \doteq \frac{1}{m} \left[1 + \frac{V(R)}{\bar{R}^2} \right] \left[\sum_{i=0}^{n} u_{t-i}^2 P_i (1-P_i)(1-c_i)^2 \right.$$

$$\left. - \sum_{i=0}^{n} \sum_{\substack{k=0 \\ k \neq i}}^{n} u_{t-i}\, u_{t-k} P_i P_k (1-c_i)(1-c_k) \right] . \qquad (8)$$

Im Vergleich zu $V(E_t)$ wird sie meistens so klein sein, daß man sie vernachlässigen kann. In einem Extremfall verschwindet sie vollständig:

Wenn die Periodenumsätze im Zeitablauf konstant sind, d. h.

$$u_{t-i} = u \ , \ i=0,\ldots,n \ ,$$

und die Verweilzeitverteilung volldeckend ist, d. h.

$$c_i = 0 \ , \ i=0,\ldots,n \ ,$$

$$\sum_{i=0}^{n} P_i = 1 \ ,$$

lautet die Varianz der Fehlerkomponente $\Sigma u_{t-i}(W_i - \hat{W}_i)$

$$V\left(\sum_{i=0}^{n} u_{t-i}\hat{W}_i\right) \doteq \frac{1}{m}\left[1 + \frac{V(R)}{\bar{R}^2}\right] u^2 \left[\sum_{i=0}^{n} P_i(1-P_i) - \sum_{i=0}^{n} \sum_{\substack{k=0 \\ k\neq i}}^{n} P_i P_k\right]$$

$$\doteq \frac{1}{m}\left[1 + \frac{V(R)}{\bar{R}^2}\right] u^2 \left[\sum_{i=0}^{n} P_i(1-P_i) - \sum_{i=0}^{n} P_i \sum_{\substack{k=0 \\ k\neq i}}^{n} P_k\right]$$

oder, wegen

$$\sum_{\substack{k=0 \\ k\neq i}}^{n} P_k = \sum_{k=0}^{n} P_k - P_i = 1 - P_i \ ,$$

$$V\left(\sum_{i=0}^{n} u_{t-i}\hat{W}_i\right) \doteq \frac{1}{m}\left[1 + \frac{V(R)}{\bar{R}^2}\right] u^2 \left[\sum_{i=0}^{n} P_i(1-P_i) - \sum_{i=0}^{n} P_i(1-P_i)\right]$$

$$\doteq 0 \ .$$

Die Varianz der betrachteten Fehlerkomponente verschwindet also bei konstanten Periodenumsätzen und volldeckendem Spektrum, da dann auch bei Verwendung eines von der Verweilzeitverteilung der Grundgesamtheit abweichenden Spektrums der Fehlerbeitrag $\Sigma u_{t-i}(W_i - \hat{W}_i)$ stets Null ist. Die Schätzfehler $(W_i - \hat{W}_i)$ gleichen sich nämlich bei der Summation aus:

$$\sum_{i=0}^{n} u(W_i - \hat{W}_i) = u \sum_{i=0}^{n} (W_i - \hat{W}_i) = 0 \ .$$

Der Fall exakt konstanter Periodenumsätze ist zwar ausgesprochen wirklichkeitsfremd, ein gewisser Ausgleich wird durch die Summation jedoch auch in weniger extremen Fällen erzielt, da die Differenzen $(W_i - \hat{W}_i)$ für einige i positiv und für einige negativ ausfallen. Sind die w_i aus einer großen Stichprobe ermittelt worden, so treten darüberhinaus nur geringe Werte für die absoluten Beträge $|W_i - \hat{W}_i|$ auf und man kann die Fehlerkomponente $\Sigma u_{t-i}(W_i - \hat{W}_i)$ dann bei praktischen Anwendungen vernachlässigen. Die Varianz des Prognosefehlers F_t läßt sich somit bei genügend großem Stichprobenumfang m und nicht allzu sprunghafter Umsatzentwicklung vielfach bereits durch die Varianz von E_t recht gut ausdrücken:

$$V(F_t) \doteq V(E_t)$$

$$\doteq \sum_{i=0}^{n} m_{t-i}(1-c_i)^2 \left[\bar{R}^2 P_i(1-P_i) + V(R)P_i\right] . \qquad (9)$$

Sie ist außer von m_{t-i} insbesondere auch vom Quadrat des mittleren Rechnungsbetrags und der Varianz der Beträge r_j abhängig, wobei jedoch der Einfluß von m_{t-i} überwiegt[1].

Als ein Ex-ante-Maß der Prognosegenauigkeit bietet sich der Standardfehler $\sqrt{V(F_t)}$ an. Er steigt ceteris paribus mit zunehmender Anzahl von Rechnungserstellungen pro Periode. Das ist plausibel, da man bei steigenden Umsätzen auch größere absolute Prognosefehler erwarten wird. Der relative Standardfehler

$$\frac{\sqrt{V(F_t)}}{\hat{e}_t} \doteq \frac{\sqrt{\sum_{i=0}^{n} m_{t-i}(1-c_i)^2\left[\bar{R}^2 P_i(1-P_i) + V(R)P_i\right]}}{\sum_{i=0}^{n} m_{t-i}\,\bar{r}_{t-i}w_i} \qquad (10)$$

---

1 Eine von EDIN durchgeführte Simulation der Einnahmen unter den genannten Annahmen erbrachte ein ähnliches Ergebnis. Vgl. R. Edin: Wirkungsweise..., a.a.O., S. 751 ff.

sinkt jedoch mit zunehmendem m_{t-i}, weil der Nenner schneller wächst als der Zähler. Hier ist besonders deutlich zu erkennen, wie wichtig große Werte von m_{t-i} für die Prognose mit der Verweilzeitverteilung sind.

Wenn auch der in (10) angegebene Zusammenhang nur für den Spezialfall der Unabhängigkeit von Rechnungsbetrag und Verweilzeit gilt, so kann man doch zumindest in der Tendenz ganz allgemein einen ähnlichen Einfluß der Größen m_{t-i}, $\overline{R}$ und $V(R)$ auf den Prognosefehler vermuten.

Ebenfalls für den Fall der Unabhängigkeit soll abschließend untersucht werden, welche Auswirkungen auf die Varianz $V(F_t)$ sich daraus ergeben, daß die Beträge $r_{j,t-i}$ der einzelnen Rechnungen nicht explizit berücksichtigt werden.

Der aus der j-ten Rechnung der Periode t-i in t eingehende Betrag $b_{ij,t-i}$ liefert im oben verwendeten Ansatz einen Varianzbeitrag von $V(B_i)$ zur Varianz $V(E_t)$. Dahinter steht die Vorstellung, daß $b_{ij,t-i}$ jede der Ausprägungen von B_i annehmen kann. Das ist tatsächlich aber nicht der Fall. Wenn keine Teilzahlungen auftreten, kann $b_{ij,t-i}$ nämlich nur den Wert $(1-c_i)r_{j,t-i}$ annehmen oder den Wert Null, je nachdem ob die betrachtete Rechnung in t bezahlt wird oder nicht. Da bei Unabhängigkeit von Betrag und Verweilzeit die Bezahlung in t mit der Wahrscheinlichkeit P_i erfolgt, lautet der Varianzbeitrag aus dieser Rechnung tatsächlich

$$\left[(1-c_i)\,r_{j,t-i}-(1-c_i)r_{j,t-i}P_i\right]^2 P_i + \left[0-(1-c_i)r_{j,t-i}P_i\right]^2(1-P_i)$$

$$= (1-c_i)^2 r_{j,t-i}^2 P_i(1-P_i)\ . \qquad (11)$$

Die Varianz von E_t ergibt sich durch Summieren der Beiträge aller beteiligten Rechnungen als

$$V(E_t) = \sum_{i=0}^{n} \sum_{j=1}^{m_{t-i}} (1-c_i)^2 r_{j,t-i}^2 P_i (1-P_i)$$

$$= \sum_{i=0}^{n} (1-c_i)^2 P_i (1-P_i) \sum_{j=1}^{m_{t-i}} r_{j,t-i}^2 \quad . \qquad (12)$$

Wenn man berücksichtigt, daß die Varianz der Beträge der in t-i erstellten Rechnungen

$$V(R_{t-i}) = \frac{1}{m_{t-i}} \sum_{j=1}^{m_{t-i}} r_{j,t-i}^2 - \bar{r}_{t-i}^2$$

lautet, wird (12) zu

$$V(E_t) = \sum_{i=0}^{n} (1-c_i)^2 P_i (1-P_i) \left[m_{t-i} V(R_{t-i}) + m_{t-i} \bar{r}_{t-i}^2 \right]$$

$$= \sum_{i=0}^{n} \left\{ m_{t-i} (1-c_i)^2 \left[\bar{r}_{t-i}^2 P_i (1-P_i) + V(R_{t-i}) P_i \right] \right\}$$

$$- \sum_{i=0}^{n} m_{t-i} (1-c_i)^2 V(R_{t-i}) P_i^2 \quad . \qquad (13)$$

Im Unterschied zum Ansatz ohne Berücksichtigung der Rechnungsbeträge $r_{j,t-i}$ (vgl.(9)) spielen hier die tatsächlich in den einzelnen Perioden t-i aufgetretenen Mittelwerte und Varianzen der Rechnungsbeträge eine Rolle. Zur Vereinfachung seien so große Stichprobenumfänge m_{t-i} unterstellt, daß man $V(R_{t-i})$ durch $V(R)$ und $\bar{r}_{t-i}$ durch $\bar{R}$ ersetzen kann. Dann wird (13) zu

$$V(E_t) \doteq \sum_{i=0}^{n} \left\{ m_{t-i} (1-c_i)^2 \left[\bar{R}^2 P_i (1-P_i) + V(R) P_i \right] \right\}$$

$$- \sum_{i=0}^{n} m_{t-i} (1-c_i)^2 V(R) P_i^2 \quad . \qquad (14)$$

Ein Vergleich mit (9) zeigt: Bei Unabhängigkeit von Rechnungsbetrag und Verweilzeit führt der Modellansatz o h n e Berücksichtigung der Rechnungsbeträge zu einer größeren Fehlervarianz $V(F_t)$ als der Ansatz m i t Berücksichtigung. Die Differenz beträgt etwa

$$V(R) \sum_{i=0}^{n} m_{t-i}(1-c_i)^2 P_i^2 \ .$$

Sie entsteht durch das unterschiedliche Ausmaß der verwendeten Information.

Ist die Unabhängigkeit tatsächlich gegeben, bereitet auch ohne Kenntnis der einzelnen Rechnungsbeträge $r_{j,t-i}$ eine Korrektur keine Schwierigkeiten. Wie groß die Differenz im allgemeineren Fall ist, der auch stochastische Abhängigkeit einbezieht, läßt sich in praktisch verwertbarer Form jedoch nicht angeben. Der Varianzbeitrag der Rechnung j aus t-i lautet dann nämlich

$$\left[(1-c_i)r_{j,t-i} - (1-c_i)r_{j,t-i}P_{i/j,t-i}\right]^2 P_{i/j,t-i}$$

$$+\left[\ 0 - (1-c_i)r_{j,t-i}P_{i/j,t-i}\ \right](1-P_{i/j,t-i})$$

$$= (1-c_i)^2 r_{j,t-i}^2 \ P_{i/j,t-i}(1-P_{i/j,t-i}) \ , \qquad (15)$$

wobei $P_{i/j,t-i}$ die - in der Regel unbekannte - individuelle Wahrscheinlichkeit dieser Rechnung dafür ist, nach i Perioden beglichen zu werden. Man erhält damit

$$V(E_t) = \sum_{i=0}^{n} \sum_{j=1}^{m_{t-i}} (1-c_i)^2 r_{j,t-i}^2 P_{i/j,t-i}(1-P_{i/j,t-i}). \qquad (16)$$

Der Vergleich mit der Varianzformel ohne Berücksichtigung der Einzelbeträge $r_{j,t-i}$ führt nun nicht mehr zu tatsächlich brauchbaren Korrekturen.

4.2.3 Intervallprognose der Einnahmen

Bei großer Anzahl von Rechnungserstellungen pro Periode läßt sich eine Intervallprognose für die tatsächlichen Einnahmen e_t als

$$\hat{e}_t - z\sqrt{V(F_t)} \leq e_t \leq \hat{e}_t + z\sqrt{V(F_t)} \qquad (1)$$

bzw., da $V(F_t)$ unbekannt ist und geschätzt werden muß, als

$$\hat{e}_t - t\sqrt{\hat{V}(F_t)} \leq e_t \leq t\sqrt{\hat{V}(F_t)} \qquad (2)$$

abgeben.

Zur Berechnung der Varianz[1]

$$\hat{V}(F_t) = \sum_{i=0}^{n} m_{t-i}\hat{V}(B_i) \qquad (3)$$

benötigt man die Größen m_{t-i} und $\hat{V}(B_i)$ (i=0,...,n). Soweit es sich bei den m_{t-i} um Werte der Vergangenheit handelt, ist ihre Ermittlung unproblematisch. Schwieriger gestaltet sich die Bestimmung zukünftiger Werte von m_{t-i}, da üblicherweise nur der Umsatz und nicht die Anzahl der Rechnungen kommender Perioden prognostiziert wird. Es wurde bereits darauf hingewiesen, daß bei großem m_{t-i} unter den geforderten Bedingungen $\bar{r}_{t-i}$ praktisch gleich $\bar{R}$ ist. Aus den prognostizierten Umsätzen $\hat{u}_{t-i}$ zukünftiger Perioden läßt sich m_{t-i} deshalb als

$$\hat{m}_{t-i} = \frac{\hat{u}_{t-i}}{\bar{R}} \qquad (4)$$

schätzen. Da $\bar{R}$ selbst unbekannt ist, muß ein Schätzwert $\hat{\bar{R}}$ verwendet werden. Man erhält $\hat{m}_{t-i}$ somit als

$$\hat{m}_{t-i} = \frac{\hat{u}_{t-i}}{\hat{\bar{R}}} \; . \qquad (5)$$

Es sei davon ausgegangen, daß die Betragsverteilung nach der Verhältnisschätzung aus einer Stichprobe von m Rechnungen ermittelt worden ist. Dann liegt mit

$$\bar{r} = \frac{1}{m}\sum_{j=1}^{m} r_j \qquad (6)$$

bereits eine Schätzung von $\bar{R}$ vor. Ferner bereitet es dann auch keine Schwierigkeiten, die Varianzen

---

1 Auf den Varianzbeitrag $V(\Sigma u_{t-i}(W_i - \hat{W}_i))$ sei hier verzichtet. Sollte er in einzelnen Fällen so groß sein, daß man ihn nicht vernachlässigen kann, so wirft seine Berücksichtigung keine besonderen Probleme auf, da keine über die Berechnung von $V(E_t)$ hinaus erforderlichen Größen benötigt werden.

$$\hat{V}(B_i) = \frac{1}{m-1}\left(\sum_{j=1}^{m} b_{ij}^2 - m\bar{b}_i^2\right), \quad i=0,\ldots,n \tag{7}$$

zu berechnen, da die Größen $\sum_j b_{ij}^2$ und $\bar{b}_i$ schon aus der Schätzung von $V(\hat{W}_i)$ und W_i bekannt sind. Der Berechnung von $V(F_t)$ steht damit nichts mehr im Wege.

Es fehlt noch die Angabe der Freiheitsgrade der Varianz $\hat{V}(F_t)$, die zur Wahl von t benötigt wird[1]. Dazu sei daran erinnert, daß der Berechnung von $\hat{V}(F_t)$ über die Linearkombination

$$\hat{V}(F_t) = \sum_{i=0}^{n} m_{t-i}\, \hat{V}(B_i)$$

eine Stichprobe von m Rechnungen zugrunde liegt. An jeder Rechnung j werden(n+1) Beobachtungen $b_{oj},\ldots,b_{nj}$ gemacht. Bei einem Stichprobenumfang von m Rechnungen werden somit m(n+1) Beobachtungen zur Schätzung von $V(F_t)$ verwendet. Bei gegebener Summe $\sum_i \sum_j b_{ij}$ ergeben sich zunächst m(n+1)-1 Freiheitsgrade. Davon entfallen aber einige aufgrund der Modellannahmen und der Bildung der Linearkombination:

- Da Teilzahlung ausgeschlossen ist, liegen für die Rechnung j sämtliche Beträge b_{ij} (i=0,...,n) fest, wenn die Verweilzeit i feststeht. Pro Rechnung entfallen auf diese Weise (n+1)-1=n Freiheitsgrade. Für alle m Rechnungen gehen m·n Freiheitsgrade verloren.

- (n+1) Freiheitsgrade entfallen, weil (n+1) Mittelwerte $\bar{b}_i$ berechnet werden müssen.

- Schließlich fällt noch ein Freiheitsgrad wegen der Linearkombination $V(F_t) = \sum m_{t-i} V(B_i)$ fort.

Es bleiben m-n-3 Freiheitsgrade erhalten. Bei großem Stichprobenumfang m sind das selbst für lange Verweilzeitverteilungen

---

1 Zur folgenden Bestimmung der Freiheitsgrade vgl. I. J. Good: What are Degrees of Freedom?, The American Statistican 27 (1973), S. 227 ff.

so viele, daß ohne großen Fehler weiterhin mit dem Faustwert $t = 2$ gearbeitet werden kann, wenn ein Signifikanzniveau von $\alpha = 0{,}05$ erwünscht ist. Damit sind alle Größen des Prognoseintervalls bestimmt.

Die Kritik der Intervallprognose gemäß (2) soll sich innerhalb der gesetzten Prämissen vollziehen. Wie wirklichkeitsnah oder -fern diese Annahmen sind, sei hier also außer Acht gelassen.

Zunächst ist festzustellen, daß eine Intervallprognose auf diese Art nur möglich ist, wenn die Varianzen $V(B_i)$ geschätzt werden können. Dazu muß die Verweilzeitverteilung mit Hilfe der "reinen" Verhältnisschätzung oder der modifizierten Verhältnisschätzung aus einer vollständigen Stichprobe ermittelt worden sein. Diese Restriktion wiegt jedoch nicht allzu schwer, da die beiden Verfahren wegen ihrer Einfachheit und ihrer geringen Anforderungen an die Organisation der Daten gegenüber den anderen Ermittlungsverfahren ohnehin die breitere Anwendungsmöglichkeit besitzen.

Ein anderer Kritikpunkt wiegt schwerer. Je weiter die Prognose in die Zukunft gerichtet ist, umso mehr Schätzwerte $\hat{m}_{t-i}$ für die Anzahl der in den einzelnen Perioden erstellten Rechnungen müssen anstelle der tatsächlichen Werte m_{t-i} verwendet werden. Die Aussage "Mit der Wahrscheinlichkeit $(1-\alpha)$ fällt e_t in das angegebene Intervall" gilt dann nur noch unter der Bedingung, daß die prognostizierten Werte $\hat{m}_{t-i}$ tatsächlich eintreten. Dadurch wird die Intervallprognose mit zunehmendem zeitlichen Abstand zwischen Abgabe- und Zielperiode immer ungenauer.

Dieser Vorwurf trifft jedoch nicht nur die Intervallprognose[1]. Er gilt in entsprechender Form auch für die Punktprognose $\hat{e}_t$,

---

1 Vgl. J. Hauschildt: Organisation der finanziellen Unternehmensführung - Eine empirische Untersuchung, Stuttgart 1970, S. 146; E. Witte u. H. Klein, a.a.O., S. 142 f.

die sich ebenfalls mit wachsender zeitlicher Distanz einer zunehmenden Anzahl von prognostizierten anstelle von bereits realisierten Inputwerten bedienen muß. So unbefriedigend es auch sein mag, einen Fehler damit zu entschuldigen, daß er an anderer Stelle ebenfalls begangen wird, der oben erhobene Vorwurf erfährt damit zumindest eine Relativierung.

4.3 Nachträgliche Beurteilung der Prognosegenauigkeit

Die Leistungsfähigkeit des Prognosemodells läßt sich durch den nachträglichen Vergleich von prognostizierten und tatsächlich hereingeflossenen Einnahmen beurteilen[1]. Üblicherweise verwendet man eine bestimmte Anzahl zeitlich nacheinander aufgetretener Prognosefehler $f_t = e_t - \hat{e}_t$ und verdichtet sie mit einem Fehlermaß. Wie weit man dazu in die Vergangenheit zurückgehen sollte, läßt sich nicht generell angeben. Werden auch weit zurückliegende Fehler berücksichtigt, so läßt sich durch die große Zahl der betrachteten Fälle die Aussagekraft der Fehlermaße unter Umständen erhöhen. Es kann aber auch genau das Gegenteil eintreten[2], weil sehr alte Werte häufig nicht mehr repräsentativ für die gegenwärtigen Verhältnisse sind und gegenläufige Entwicklungen sich in ihrem Einfluß auf einige Fehlermaße aufheben.

Vor der Berechnung von Fehlermaßen muß auch geklärt werden, welcher Prognosewert $\hat{e}_t$ überhaupt in den Fehler $f_t = e_t - \hat{e}_t$ eingehen soll. Nachträglich stehen nämlich mehrere Werte für $\hat{e}_t$ zur Auswahl, die zu verschiedenen Zeitpunkten vor der Periode t ermittelt worden sind. Die Ursache dafür liegt im Verfahren selbst. In jeder Periode t wird eine Prognose für alle

---

1 Vgl. J. Schwarze, a.a.O., S. 551.

2 Vgl. H. D. Haustein: Prognoseverfahren in der sozialistischen Wirtschaft, Berlin 1970, S. 274.

zukünftigen Perioden bis einschließlich t+n abgegeben, die Einnahmen einer Periode werden somit insgesamt n-mal vorhergesagt. Die Fehlerbetrachtung kann sich nun auf einen beliebigen dieser n Prognosewerte beziehen. Wichtig ist nur, daß alle zur Berechnung des Fehlermaßes verwendeten Abweichungen einheitlich von Prognosewerten desselben zeitlichen Abstandes zwischen Abgabe- und Zielperiode der Prognose ausgehen. Welcher Abstand gewählt wird, richtet sich nach der konkreten Fragestellung. Große Abstände lassen auch die Fehler, die bei der Vorhersage der Umsätze gemacht wurden, in die Beurteilung eingehen, während extrem kleine Abstände zu exakteren Inputgrößen führen[1] und insbesondere dann verwendet werden können, wenn das Ausmaß der Übereinstimmung von Prognosemodell und Wirklichkeit ermittelt werden soll. Da die Einnahmeprognosen gewöhnlich umso ungenauer werden, je weiter sie in die Zukunft gerichtet sind, und der zeitliche Abstand auf diese Weise den Wert eines Fehlermaßes wesentlich beeinflussen kann, sollte er bei Vergleichen mit anderen Prognoseergebnissen stets angegeben werden.

Die in der Literatur vorgeschlagenen Maße zur Erfassung des Prognosefehlers lassen sich in zwei Kategorien einteilen. Einmal wird untersucht, ob die Wendepunkte in der Zeitreihe der Prognosewerte richtig vorhergesagt worden sind, zum anderen richtet sich das Interesse auf den numerischen Wert der Abweichung der Realisation von der Prognose[2]. Im Fall der Einnahmeprognose steht die zweite Betrachtung im Vordergrund. Man will wissen, um welchen Betrag man sich bei der Abgabe der Prognose verschätzt hat. Bei Überschätzungen der tatsächlichen Einnahmen können gefährliche Liquiditätsengpässe entstehen, bei Unterschätzung Gewinneinbußen wegen unnötiger Kassenhaltung[3].

---

1 Vgl. R. Lewandowski: Prognosesysteme und ihre heutigen Anwendungsmöglichkeiten in der Unternehmensführung, in: G. Bamberg u. O. Opitz: Information und Prognose, Meisenheim a. Glan 1975, S. 61.

2 Vgl. K. W. Rothschild: Zur Frage der Erfolgsbeurteilung ökonomischer Prognosen, ZgesStw 130 (1974), S. 580.

3 Vgl. E. Witte u. H. Klein, a.a.O., S. 145.

Optimal wäre ein Fehlermaß, das die Abweichungen gemäß ihren Auswirkungen gewichtete[1]. Solch eine Verlustfunktion ist für den Fehler der Einnahmeprognose aber nur sehr schwer vorstellbar, da seine Auswirkungen selten quantifizierbar sein dürften. Die Ansprüche müssen deshalb reduziert werden.

Ein Fehlermaß sollte zumindest Vergleiche mit Ergebnissen früherer Untersuchungen innerhalb desselben Unternehmens zulassen. Darüberhinaus ist Vergleichbarkeit mit anderen Unternehmen wünschenswert. Besonders günstig ist es, wenn der Wert des Maßes bereits ohne Vergleich eine absolute Aussage über die Güte der Prognose erlaubt. Im folgenden werden ohne Anspruch auf Vollständigkeit einige Maße beschrieben, die diesen Anforderungen in unterschiedlichem Umfang genügen. Eine Bewertung läßt sich daraus jedoch nicht ableiten, da die Aussagen zu verschiedenartig sind.

Ein einfaches Maß für die Überprüfung der Erwartungstreue der Prognose ist die mittlere Abweichung

$$MA = \frac{1}{T} \sum_{t=1}^{T} f_t \quad . \tag{1}$$

Sie weicht bei Unverzerrtheit nicht wesentlich von Null ab[2]. LANGEN schlägt vor, aus MA und der Standardabweichung der betrachteten T Fehler

$$s_{f_t} = \sqrt{\frac{1}{T} \sum_{t=1}^{T} (f_t - MA)^2}$$

die Größe

$$t = \frac{MA \cdot \sqrt{T-1}}{s_{f_t}} \tag{2}$$

---

1 Vgl. J. Schwarze, a.a.O., S. 540.

2 Vgl. R. Edin: Wirkungsweise..., a.a.O., S. 750.

zu berechnen und mit dem einseitigen t-Test zu prüfen, ob MA signifikant von Null verschieden ist[1]. Als eine Faustregel hat diese Prüfung ihre Berechtigung, ein echter Signifikanztest ist sie indes nicht, da die Bedingungen des t-Tests nicht erfüllt sind[2]. Insbesondere sind die Beobachtungen f_t nicht voneinander unabhängig; da bei diesem Prognoseverfahren die Überschätzung von Einnahmen einer Periode tendenziell die Unterschätzung von Einnahmen späterer Perioden (und umgekehrt) bewirkt[3], ist die Zeitreihe der Prognoseabweichungen negativ autokorreliert.

Um die Vergleichbarkeit zu erhöhen, kann man die Abweichungen relativieren, indem man sie durch $\hat{e}_t$ dividiert[4]. Man erhält dann die mittlere relative Abweichung

$$MA_r = \frac{1}{T} \sum_{t=1}^{T} \frac{f_t}{\hat{e}_t} , \qquad (3)$$

die bei Unverzerrtheit der Prognosen ebenfalls nicht wesentlich von Null verschieden ist.

Eine andere Möglichkeit, Verzerrungen zu erkennen, beruht auf folgender Überlegung[5]: In der Abbildung 11 gibt die 45°-Linie die perfekten Prognosen an. Hier gilt

$$e_t = \hat{e}_t .$$

---

1 Vgl. H. Langen et al., a.a.O., S. 68 ff.

2 Vgl. zum t-Test z. B. S. Siegel: Nonparametric Statistics for the Behavioral Sciences, Tokyo 1956, S. 19.

3 Vgl. H. J. Schmitt: Die Planrevision..., a.a.O., S. 26.

4 Vgl. R. Edin: Wirkungsweise..., a.a.O., S. 750.

5 Vgl. H. Theil: Economic Forecasts and Policy, Second revised edition, Amsterdam 1965, S. 31 ff; ders.: Applied Economic Forecasting, Amsterdam 1966, S. 33 ff.

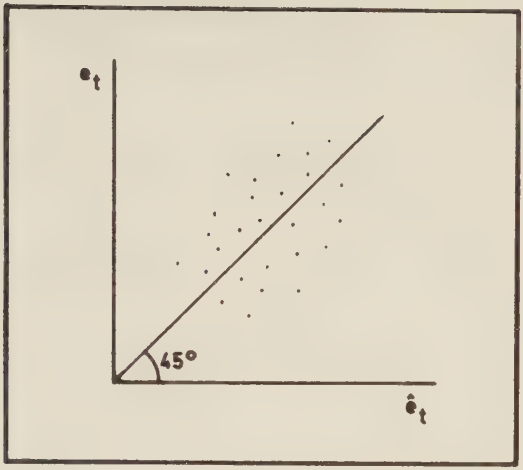

Abb. 11: Regression von e_t auf $\hat{e}_t$.
Quelle: J. Schwarze, a.a.O., S. 552.

Ergibt die Regression

$$e_t = a + b\hat{e}_t \qquad (4)$$

a=0 und b=1, so fällt sie mit der 45^o-Linie zusammen und zeigt Erwartungstreue der Prognose an. Wenn a und b von diesen Werten abweichen, erfordert es ein gewisses Maß an Fingerspitzengefühl zu beurteilen, ob diese Abweichungen zufällig oder systematisch verursacht sind. Die Aussagekraft von entsprechenden Signifikanztests ist aus dem bereits genannten Grund jedenfalls eingeschränkt.

Zur Beurteilung sollte man zusätzlich den Produktmomentkorrelationskoeffizienten $r_{e_t\hat{e}_t}$ heranziehen. Sein absoluter Wert läßt zwar ohne Signifikanztest nur vage Aussagen über die Korrelation von e_t und $\hat{e}_t$ zu[1], je näher jedoch sein Quadrat $r^2_{e_t\hat{e}_t}$

1 Vgl. G. Clauß u. H. Ebner: Grundlagen der Statistik für Psychologen, Pädagogen und Soziologen, 2., neubearbeitete und erweiterte Auflage, Zürich/Frankfurt a.M. 1975, S. 122.

bei 1 liegt, umso stärker ist, unabhängig von der Verteilung der e_t und $\hat{e}_t$, der Grad der Linearität[1], und umso eher wird man bereit sein, auch kleinere Abweichungen von der 45^o-Linie als überzufällig anzusehen.

Die Berechnung von a und b ist relativ aufwendig, gemessen an der von MA und MA_r. Sie bietet jedoch die Möglichkeit, einmal aufgedeckte systematische Fehler mit Hilfe der Gleichung (4) bei zukünftigen Prognosen zu berücksichtigen[2]. Es muß noch angemerkt werden, daß die Regression $e_t = a + b\ \hat{e}_t$ natürlich nur berechnet werden kann, wenn die betrachteten $\hat{e}_t$ überhaupt variieren.

Die Güte der Anpassung der Zeitreihe der prognostizierten Einnahmen an die der tatsächlichen Einnahmen läßt sich durch die mittlere Abweichung ohne Berücksichtigung des Vorzeichens ausdrücken[3]:

$$MA = \frac{1}{T} \sum_{t=1}^{T} |f_t| \quad . \tag{5}$$

Die Vergleichbarkeit wird auch hier durch Division durch $\hat{e}_t$ erhöht[4]. Die mittlere relative Abweichung ohne Berücksichtigung des Vorzeichens gibt an, wieviel Prozent des jeweiligen prognostizierten Wertes $\hat{e}_t$ die Abweichung im Durchschnitt beträgt[5]:

$$MA_r = \frac{1}{M} \sum_{t=1}^{T} \frac{|f_t|}{\hat{e}_t} \quad . \tag{6}$$

---

1 Vgl. R. M. Meyer: On Correlation and Linearity, The American Statistician 28 (1974), S. 103.

2 Vgl. H. Theil: Applied Economic Forecasting, a.a.O., S. 34.

3 Vgl. R. Edin: Wirkungsweise..., a.a.O., S. 750.

4 Auch Division durch e_t wird vorgeschlagen, vgl. J. Schwarze, a.a.O., S. 556.

5 Vgl. E. Witte u. H. Klein, a.a.O., S. 144.

Den gleichen Zweck wie MA und MA_r erfüllen die Wurzel der mittleren quadratischen Abweichung[1]

$$MQA = \frac{1}{T} \sum_{t=1}^{T} f_t^2 \qquad (7)$$

und die Wurzel der mittleren relativen quadratischen Abweichung[2]

$$MQA_r = \frac{1}{T} \sum_{t=1}^{T} \left(\frac{f_t}{\hat{e}_t}\right)^2 .$$

$\sqrt{MQA}$ und $\sqrt{MQA_r}$ sprechen besonders stark auf Ausreißer an, können aber nicht mehr unmittelbar interpretiert werden.

1 Vgl. R. Edin: Wirkungsweise..., a.a.O., S. 750.

2 Ebenda.

5 Überwachung der Stabilität der Verweilzeitverteilung

5.1 Vorbemerkungen

Eine wesentliche Annahme der Einnahmeprognose mit Hilfe von Verweilzeitverteilungen ist die der Stabilität des verwendeten Liquidationsspektrums im Zeitablauf. Im Rahmen des stochastischen Prognoseansatzes liegt Stabilität dann vor, wenn die jeweiligen Verteilungen, nach denen sich die Umsätze einzelner Perioden tatsächlich liquidieren (und die deshalb Periodenspektren genannt werden sollen), nur in der Größenordnung von zufälligen Abweichungen von einer mittleren Verweilzeitverteilung verschieden sind[1], die mit derjenigen der fiktiven Grundgesamtheit identisch ist. Die Stabilität der Verweilzeitverteilung äußert sich mit anderen Worten darin, daß die Periodenspektren als Stichprobenrealisation der Verteilung der Grundgesamtheit angesehen werden können.

Ein Stabilitätstest, der an diese Definition anknüpft, hat demnach die Frage zu beantworten, ob man annehmen darf, daß zwei zu verschiedenen Zeitpunkten ermittelte Liquidationsspektren aus derselben Grundgesamtheit stammen. Diese Fragestellung entspricht der eines Homogenitätstests.

Geprüft wird auf diese Weise allerdings nur, ob eine unbekannte stabile Verteilung vorliegt, während den praktischen Anwender etwas ganz anderes interessiert. Er hat bisher ein bestimmtes Liquidationsspektrum zur Prognose verwendet und möchte wissen: Darf ich dieses Spektrum weiterhin verwenden, oder gibt es das augenblickliche Zahlungsverhalten der Kunden so schlecht wieder, daß ich es neu ermitteln muß[2]? Es wird somit nicht die

---

1 Vgl. R. Edin: Dynamische Analyse..., a.a.O., S. 59; H. Kossbiel, a.a.O., S. 194.

2 Vgl. auch L. Neumeyer, a.a.O., S. 72.

Stabilität eines unbekannten Spektrums einer fiktiven Grundgesamtheit angesprochen, sondern die eines tatsächlich vorliegenden. Entsprechend muß auch für die Stabilitätsüberwachung eine andere Definition der Stabilität gefunden werden. Stabilität soll im folgenden dann gegeben sein, wenn sich die Periodenspektren von dem in der Vergangenheit ermittelten Liquidationsspektrum nur zufallsbedingt unterscheiden. Ein entsprechender Stabilitätstest hat somit zu prüfen, ob angenommen werden darf, daß zur Kontrolle entnommene Stichproben einer Grundgesamtheit entstammen, in der die bisher verwendete Verweilzeitverteilung gilt. Gesucht ist demnach ein geeigneter Anpassungstest.

Es sei hier nur am Rande darauf hingewiesen, daß diese Definition der Stabilität gegenüber der weiter oben gegebenen nicht nur größere Praxisbezogenheit aufweist, sondern darüber hinaus Veränderungen des Spektrums im Zeitablauf auch eher erkennen läßt. Wurde z. B. bislang mit dem Spektrum $(w_0\ w_1 \ldots w_n)$ gearbeitet und lautet ein zur Kontrolle neu ermitteltes Spektrum $(w_0^+\ w_1^+ \ldots w_n^+)$, so wird nach beiden Definitionen geprüft, ob $w_i^+ - w_i$ für alle i nur zufällig von Null verschieden ist. Dabei ist w_i^+ die Realisation einer Zufallsgröße, die mit W_i^+ bezeichnet werden soll. Je nach der verwendeten Modellannahme ist w_i nun entweder die Ausprägung einer Zufallsvariablen[1] W_i oder eine vorgegebene Konstante, und damit ergeben sich unterschiedliche Varianzen für die betrachtete Differenz. Ist gefragt, ob beide Stichproben aus einer Grundgesamtheit stammen, so ist bei Unabhängigkeit der beiden Stichproben[2]

$$V(W_i^+ - W_i) = V(W_i^+) + V(W_i) \ .$$

---

1 W_i ist identisch mit der Variablen $\hat{W}_i$. Die Bezeichnung $\hat{W}_i$ ist hier unzweckmäßig, da nicht die Schätzung von $(W_0\ W_1 \ldots W_n)$ betrachtet, sondern die Möglichkeit der Verwendung von $(w_0\ w_1 \ldots w_n)$ überprüft wird.

2 Vgl. auch L. Kish u. I. Hess: On Variances of Ratios and their Differences in Multi-Stage Samples, Journal of the American Statistical Association 54 (1959), S. 423.

Bei der Untersuchung, ob die Kontrollstichprobe aus einer Grundgesamtheit mit der Verteilung $(w_0\ w_1 \ldots w_n)$ stammt, gilt dagegen

$$V(W_i^+ - w_i) = V(W_i^+) \ .$$

Hier können Abweichungen bereits als signifikant gelten, die beim Homogenitätstest noch als zufallsbedingt angesehen werden. Ob dieser Vorteil tatsächlich genutzt wird, hängt allerdings von der Wahl des angewendeten Testverfahrens ab.

Im weiteren Verlauf wird zunächst der Vergleich von zwei Verweilzeitverteilungen behandelt, der auch als direkte Stabilitätsüberwachung bezeichnet werden soll. Die im Anschluß daran vorgestellte indirekte Stabilitätsüberwachung, die auf die Überlegungen zur Intervallprognose der Einnahmen zurückgreift, setzt nicht direkt an den Liquidationsspektren an, sonder versucht, mangelnde Stabilität an der Entwicklung des Prognosefehlers zu erkennen.

Es ist noch anzumerken, daß die Überlegungen zur Stabilitätsüberwachung ausschließlich auf die Betragsverteilung abstellen. Das soll nicht nur mit der untergeordneten Bedeutung der Anzahlverteilung für die Prognose der Einnahmen gerechtfertigt werden, sondern auch damit, daß diese theoretisch erheblich einfacher zu behandeln ist als die Betragsverteilung und bei der Stabilitätskontrolle keine zusätzlichen Gesichtspunkte aufweist.

5.2 Direkte Stabilitätsüberwachung

5.2.1 Überwachung mit Hilfe der Restforderungen

In Anlehnung an einen Vorschlag von BENISHAY[1] läßt sich die Stabilität der Verweilzeitverteilung ohne großen rechnerischen Auf-

1 Vgl. H. Benishay: Managerial Controls of Accounts Receivable: A Reply, Journal of Accounting Research 4 (1966), S. 132.

wand auf folgende Weise prüfen:

Das in der Vergangenheit ermittelte Liquidationsspektrum sei $\mathbf{w} = (w_0\ w_1\ \ldots\ w_n)$. Daraus wird das kumulierte Spektrum

$$\mathbf{k} = (k_0\ k_1\ \ldots k_n) = \left(\sum_{j=0}^{0} w_j \quad \sum_{j=0}^{1} w_j \ldots \sum_{j=0}^{n} w_j\right)$$

gebildet, in dem die Komponente k_i angibt, welcher Anteil eines Periodenumsatzes erwartungsgemäß bis zur i-ten Periode nach der Fakturierung liquidiert worden ist.

Um zu prüfen, ob die Verweilzeitverteilung der Vergangenheit in einer späteren Periode t noch zutrifft, stellt man fest,

- wie hoch die Umsätze der Perioden t-n bis t waren,
- wie hoch die Forderungen sind, die in t aus diesen Umsätzen jeweils noch bestehen,

und ermittelt daraus eine neue kumulierte Betragsverteilung

$$\mathbf{k}^+ = (k_0^+\ k_1^+\ \ldots\ k_n^+)\ ,$$

in der man k_i^+ als

$$k_i^+ = \frac{\text{Umsatz in t-i} - \text{in t noch bestehende Restforderung aus Umsatz in t-i}}{\text{Umsatz in t-i}}$$

erhält. Die beiden kumulierten Verteilungen $\mathbf{k}$ und $\mathbf{k}^+$ werden dann verglichen[1].

Das sei an einem Beispiel verdeutlicht. Das Spektrum der Vergangenheit sei $\mathbf{w}$ = (0,03 0,08 0,25 0,20 0,12 0,10 0,05 0,06

---

1 Nach BENISHAYs Vorschlag werden nicht die bereits beglichenen Anteile verglichen, sondern die noch ausstehenden; er kumuliert die Verteilung $\mathbf{w}$ dazu "von rechts". Dieses Vorgehen erscheint nicht zweckmäßig, da es (wie man sich anhand des Beispiels schnell klarmachen kann) nur für volldeckende Spektren geeignet ist.

0,05 0,04). Kumulieren führt zu $\mathbf{k}$ =(0,03 0,11 0,36 0,56 0,68 0,78 0,83 0,89 0,94 0,98). Die in t ermittelte kumulierte Verteilung laute $\mathbf{k}^+$ = (0,00 0,05 0,22 0,57 0,70 0,77 0,87 0,92 0,93 0,95). Die neue Betragsverteilung $\mathbf{w}^+$ läßt sich daraus als $\mathbf{w}^+$ = (0,00 0,05 0,17 0,35 0,13 0,07 0,10 0,05 0,01 0,02) berechnen. In Abbildung 12 und 13 sind die beiden Verteilungen bzw. kumulierten Verteilungen abgetragen.

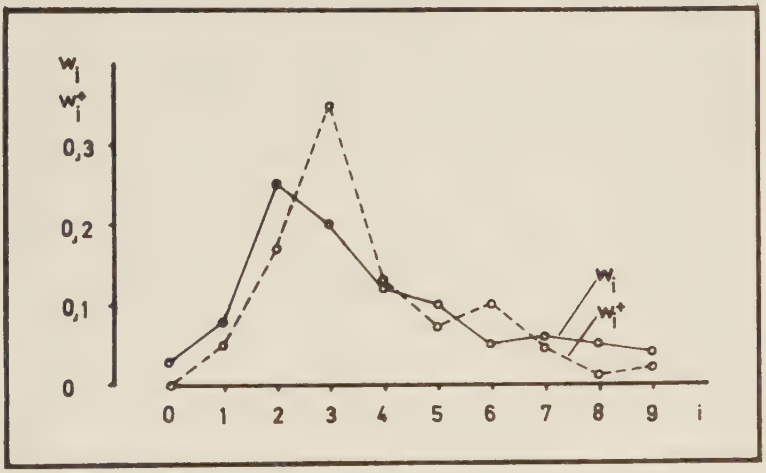

Abb. 12: Beispiel für zwei zu verschiedenen Zeitpunkten ermittelte Betragsverteilungen

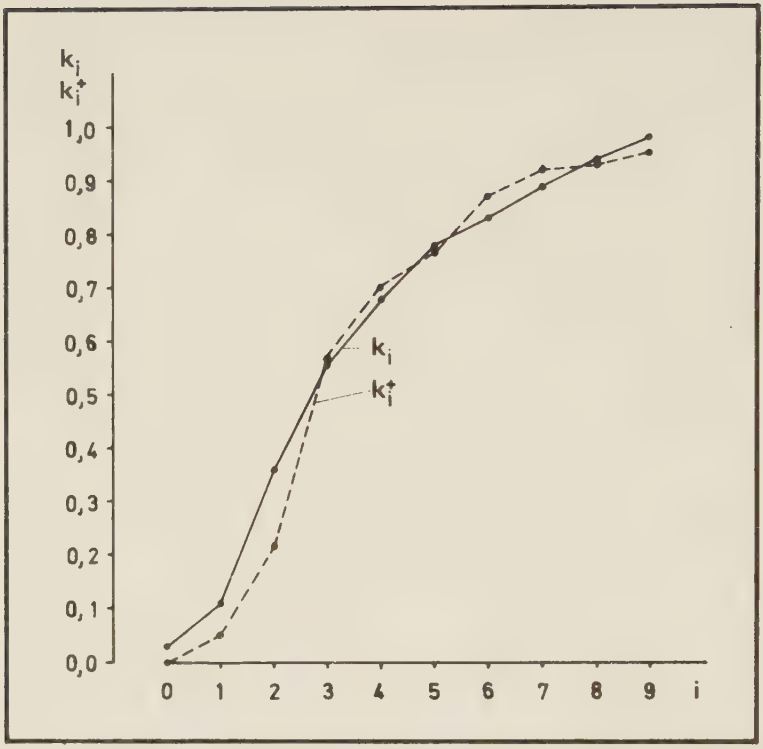

Abb. 13: Kumulierte Verteilungen der Abb. 12

Ob sich die neue Betragsverteilung signifikant von der alten unterscheidet, läßt sich nach BENISHAY mit dem Vorzeichentest und dem run-Test prüfen[1].

Dazu ermittelt man die Vorzeichen der Differenzen $k_i - k_i^+$:

1 Vgl. H. Benishay: Managerial Controls...: A Reply, a.a.O., S. 132.

i	0	1	2	3	4	5	6	7	8	9
k_i	0,03	0,11	0,36	0,56	0,68	0,78	0,83	0,89	0,94	0,98
k_i^+	0,00	0,05	0,22	0,57	0,70	0,77	0,87	0,92	0,93	0,95
Vorzeichen von $k_i - k_i^+$	+	+	+	-	-	+	-	-	+	+

Tabelle 9: Arbeitstabelle zum Vorzeichen- und run-Test

Es sei unterstellt, daß die Liquidation von Umsätzen verschiedener Perioden unabhängig erfolgt. Dann sind die Glieder der Vorzeichenfolge unabhängige Beobachtungen und lassen die Anwendung der angegebenen Tests zu[1].

Wenn das neue Spektrum nur zufällig vom alten abweicht, kann man mit einer etwa gleichgroßen Anzahl von Plus- und Minuszeichen rechnen. Ob ein signifikantes Ungleichgewicht vorliegt, wird mit dem Vorzeichentest festgestellt[2]. Im Beispiel trifft das nicht zu. Eine signifikant ungleiche Anzahl ist bei 10 Zeichen auf einem Signifikanzniveau von $\alpha = 0{,}05$ erst dann gegeben, wenn weniger als 3 Zeichen einer Art vorhanden sind[3].

Neben der Anzahl der Plus- und Minuszeichen ist ihre Reihenfolge von Bedeutung. Teilfolgen, die nur aus Zeichen einer Art bestehen, werden als run bezeichnet[4]. Im Beispiel treten 5 runs auf:

---

1 Im Beispiel ist $k_i - k_i^+$ stets von Null verschieden. Das braucht nicht immer der Fall zu sein, wenngleich $k_i - k_i^+ = 0$ bei der besonderen Art der Ermittlung von k_i und k_i^+ ausgesprochen unwahrscheinlich ist. Treten in der Vorzeichenfolge tatsächlich Nullen auf, so ist bei der Durchführung der Tests eine besondere Berücksichtigung erforderlich, auf die jedoch nicht eingegangen werden soll.

2 Vgl. L. Sachs: Statistische Auswertungsmethoden, Zweite, neubearbeitete und erweiterte Auflage, Berlin/Heidelberg/New York 1969, S. 315.

3 Ebenda, S. 316.

4 Vgl. S. Siegel, a.a.O., S. 52.

+ + +	- -	+	- -	+ +
1. run	2. run	3. run	4. run	5. run .

Wenn sich die beiden Spektren nur zufällig unterscheiden, wird man weder eine extrem kleine Anzahl langer runs, etwa + + + + + + - - - - , noch eine extrem große Anzahl kurzer runs, etwa + - + - + - + - + + , erwarten. Da bei 6 Plus- und 4 Minuszeichen erst bei weniger als 3 oder mehr als 8 runs auf signifikante Abweichung von einer zufälligen Anordnung geschlossen werden kann ($\alpha = 0{,}05$)[1], darf die Annahme, daß die Betragsverteilung des Beispiels stabil geblieben ist, nicht abgelehnt werden.

Der Stabilitätstest nach BENISHAY setzt eine geeignete Organisation der benötigten Daten voraus, die nicht in jedem Fall vorzuliegen braucht. Man kann zwar davon ausgehen, daß die Periodenumsätze bekannt sind, weil sie zur Prognose der Einnahmen benötigt werden, Schwierigkeiten können sich unter Umständen jedoch bei der Ermittlung der Restforderungen ergeben; da das betriebswirtschaftliche Rechnungswesen primär auf die Erfolgsermittlung ausgerichtet ist, sind zur Gewinnung von Daten für die Finanzplanung häufig Sonderrechnungen erforderlich[2].

Ist eine entsprechende Datenorganisation gegeben, so sind diese Tests ausgesprochen einfach durchzuführen, da sie keine eigene Stichprobenerhebung benötigen[3]. Ferner kommen sie mit einem Mindestmaß an Annahmen aus.

---

1 Vgl. S. Siegel, a.a.O., S. 252 f.

2 Vgl. E. Witte u. H. Klein, a.a.O., S. 141.

3 Es sei darauf aufmerksam gemacht, daß sich dann auch eine zusätzliche eigenständige Möglichkeit zur Ermittlung der Verweilzeitverteilung bietet, die ohne Stichprobenentnahme auskommt und aktuelle Verteilungen liefert. Eine Intervallschätzung für die einzelnen Komponenten ist bei dieser Schätzmethode allerdings nicht möglich, weil die dazu benötigten Varianzen $\hat{V}(W_i)$ nicht berechnet werden können.

Diesen Vorzügen stehen aber erhebliche Schwächen gegenüber. Bei besonders kurzen Verweilzeitverteilungen ist überhaupt keine Stabilitätskontrolle möglich, weil zur Durchführung des Vorzeichentests mindestens 5, zu der des run-Tests mindestens 8, in manchen Situationen sogar mindestens 13 Zeichen vorliegen müssen[1]. Ferner verschenken beide Tests viel von der in den kumulierten Spektren enthaltenen Information dadurch, daß sie nur die Vorzeichen und nicht die Beträge der Differenzen verwerten. Sie sind ausgesprochen konservativ. Schließlich berücksichtigen sie nicht die Anzahl der Rechnungen, die hinter dem kumulierten Spektrum k^+ stehen. Solange die Verweilzeitverteilungen auf die oben geschilderte Weise aus den Restforderungen hergeleitet werden, bringt eine explizite Berücksichtigung allerdings große Schwierigkeiten mit sich, da diese Forderungen aus verschiedenen Perioden mit unterschiedlicher Anzahl von Fakturen stammen.

Soll bei der Kontrolle auch beachtet werden, aus wieviel Rechnungen das Spektrum k^+ bzw. w^+ ermittelt worden ist, so empfiehlt es sich, die Stabilitätsüberwachung mit Hilfe eigens zu diesem Zweck gezogener Stichproben vorzunehmen. Gegenüber der Überwachung nach BENISHAY bedeutet das unter Umständen zwar einen größeren Arbeitsaufwand, dafür ist jedoch der Einsatz trennschärferer Tests möglich. In vielen Fällen wird ferner die Entnahme einer Stichprobe von Rechnungen technisch einfacher durchzuführen sein als die Feststellung der Restforderungen aus bestimmten Periodenumsätzen.

5.2.2 Überwachung durch Stichprobenentnahme

Im folgenden wird davon ausgegangen, daß zur Kontrolle eine Stichprobe von m^+ Rechnungen gezogen worden sei, in der sich

---

1 Vgl. L. Sachs, a.a.O., S. 316; S. Siegel, a.a.O., S. 252 f.

die nach der Verhältnisschätzung berechnete Betragsverteilung $\mathbf{w}^+ = (w_0^+ \; w_1^+ \; \ldots \; w_n^+)$ ergeben habe. Das bisher verwendete Liquidationsspektrum sei wieder mit $\mathbf{w} = (w_0 \; w_1 \; \ldots \; w_n)$ bezeichnet. Bei der Wahl eines geeigneten Testverfahrens spielt der Stichprobenumfang m^+ eine wichtige Rolle, da er die Verteilung der Stichprobenparameter entscheidend beeinflußt. Bei großem m^+ ist W_i^+ annähernd normalverteilt[1], und man kann einen parametrischen Test anwenden. Ist m^+ zu klein, um eine ausreichende Approximation zu bewirken, so muß ein entsprechendes nichtparametrisches Verfahren herangezogen werden. Einen verteilungsfreien Test, welcher der spezifischen Eigenschaft der Betragsverteilung vollkommen gerecht wird, gibt es allerdings nicht. Deshalb soll auf ein Verfahren zurückgegriffen werden, das den Sachverhalt wenigstens ungefähr trifft.

Es sei noch einmal an die Abbildung 13 erinnert, in der die beiden kumulierten Betragsverteilungen $\mathbf{k}$ und $\mathbf{k}^+$ verglichen werden. $\mathbf{k}^+$ sei nun das Ergebnis einer Stichprobe von m^+ Rechnungen. Stammt diese Stichprobe aus der Grundgesamtheit, in der $\mathbf{k}$ gilt, so werden sich $\mathbf{k}$ und $\mathbf{k}^+$ nicht allzu stark unterscheiden. Man wird dann - besonders bei größerem m^+ - keine großen absoluten Werte für die Differenzen $k_i - k_i^+$ erwarten. Zur Feststellung, ob die beobachteten Differenzen noch als zufällig betrachtet werden dürfen, bietet sich der KOLMOGOROFF-SMIRNOV-Anpassungstest[2] an. Er läuft folgendermaßen ab:

Man kumuliert die beiden Spektren $\mathbf{w}$ und $\mathbf{w}^+$ zu $\mathbf{k}$ bzw. $\mathbf{k}^+$, wobei wieder

$$k_i = \sum_{j=0}^{i} w_j \qquad (1)$$

---

1 Vgl. die Faustregel auf S. 49 f.

2 Vgl. dazu etwa G. A. Lienert: Verteilungsfreie Methoden in der Biostatistik, Zweite völlig neu bearbeitete Auflage, Band I, Meisenheim am Glan 1973, S. 459 ff; S. Siegel, a.a.O., S. 47 ff.

und

$$k_i^+ = \sum_{j=0}^{i} w_j^+ \qquad (2)$$

ist, und berechnet die Differenzen

$$d_i = k_i - k_i^+ \qquad (3)$$

für alle i. Der dabei beobachtete größte absolute Betrag von d_i,

$$D = \max |d_i| \quad , \qquad (4)$$

wird mit dem kritischen Wert D_c verglichen, der bei zweiseitiger Fragestellung und einem Signifikanzniveau von $\alpha = 0{,}05$ als

$$D_c = \frac{1{,}36}{\sqrt{m^+}} \qquad (5)$$

berechnet wird, wobei m^+ größer als 35 sein sollte[1]. Die Annahme der Stabilität darf nicht verworfen werden, solange D kleiner als D_c ist.

Im obigen Beispiel (es sei nun angenommen, daß $\mathbf{w}^+$ und damit auch $\mathbf{k}^+$ aus einer Stichprobe von $m^+ = 100$ Rechnungen ermittelt worden ist) ergeben sich die in Tabelle 10 angegebenen Differenzen d_i.

i	0	1	2	3	4	5	6	7	8	9
k_i	0,03	0,11	0,36	0,56	0,68	0,78	0,83	0,89	0,94	0,98
k_i^+	0,00	0,05	0,22	0,57	0,70	0,77	0,87	0,92	0,93	0,95
d_i	0,03	0,06	0,14	-0,01	-0,02	0,01	-0,04	-0,03	0,01	0,03

Tabelle 10: Arbeitstabelle zum KOLMOGOROFF-SMIRNOV-Test

1 Vgl. S. Siegel, a.a.O., S. 251.

Der größte absolute Betrag von d_i ist D = 0,14. Der kritische Wert beträgt $D_c = 1{,}36/\sqrt{100} = 0{,}136$. Die Annahme der Stabilität wird verworfen.

Der KOLMOGOROFF-SMIRNOV-Test gilt als der wirksamste der derzeit bekannten verteilungsfreien Anpassungstests[1]. Er entspricht dem Sachverhalt jedoch nur ungefähr. Da er stetige Verteilungen vergleicht[2], müßten eigentlich unendlich viele Ausprägungen der diskreten Variablen i vorhanden sein[3], das Spektrum müßte also unendlich lang sein. Es hat sich aber gezeigt, daß der KOLMOGOROFF-SMIRNOV-Test auch für eine endliche und sogar recht kleine Anzahl von Klassen i durchaus anwendbar ist[4].

Es existiert hingegen noch eine andere, schwerwiegendere Abweichung von den Modellannahmen. Das kumulierte Spektrum $\mathbf{k}^+$ kann zwar als Verteilungsfunktion aufgefaßt werden, die Komponente k_i^+ gibt jedoch an, welcher Anteil der in der Stichprobe erfaßten DM-Stücke eine Verweilzeit von i oder weniger Perioden aufweist, während der Test die Verteilungsfunktion $\mathbf{k}^+$ so behandelt, als wäre k_i^+ der Anteil der Rechnungen mit einer Verweilzeit von i oder weniger Perioden. Tatsächlich wird also gar nicht getestet, ob sich zwei B e t r a g s v e r t e i l u n g e n zufällig unterscheiden, sondern man prüft, ob ein signifikanter Unterschied vorläge, wenn $\mathbf{w}$ und $\mathbf{w}^+$ A n z a h l v e r t e i l u n g e n wären.

Die Aussagekraft des Testergebnisses wird dadurch eingeschränkt. Vermutlich ist die Wahrscheinlichkeit, die richtige Nullhypothese "Stabilität liegt vor" abzulehnen, größer als der verwen-

---

1 Vgl. G. A. Lienert, a.a.O., S. 459; S. Siegel, a.a.O., S. 51.

2 Vgl. E. Kreyszig, a.a.O., S. 234.

3 Vgl. L. Sachs, a.a.O., S. 328.

4 Vgl. M. J. Slakter: A Comparison of the Pearson Chi-Square and Kolmogorov Goodness-Of-Fit Tests with Respect to Validity, Journal of the American Statistical Association 60 (1965), S. 857.

wendete Wert von α angibt: Man stelle sich vor, aus einer Grundgesamtheit mit der Anzahlverteilung $\mathbf{p} = (p_0\ p_1\ \ldots\ p_n)$ und der Betragsverteilung $\mathbf{w} = (w_0\ w_1\ \ldots\ w_n)$ werden mehrere Stichproben vom Umfang m^+ gezogen und man ermittelt jeweils die Anzahlverteilung $\mathbf{p}^+$ und die Betragsverteilung $\mathbf{w}^+$, deren Komponenten $p_0^+, p_1^+ \ldots p_n^+$ bzw. $w_0^+, w_1^+ \ldots w_n^+$ wieder als Realisation von $P_0^+, P_1^+ \ldots P_n^+$ bzw. $W_0^+, W_1^+ \ldots W_n^+$ aufgefaßt seien. Bei einem Vergleich der Varianzen $V(P_i^+)$ und $V(W_i^+)$ wird man feststellen, daß die Komponenten der Betragsverteilung über die Stichproben hinweg tendenziell stärker variieren als die entsprechenden Komponenten der Anzahlverteilung. Für den Sonderfall der Unabhängigkeit von Beträgen und Verweilzeiten der Rechnungen wird das besonders deutlich. Wie schon gezeigt wurde[1], ist dann (bei kleinem Auswahlsatz)

$$V(W_i^+) \doteq \frac{p_i(1-p_i)}{m^+}\,(1-c_i)^2\left(1 + \frac{V(R)}{\bar{R}^2}\right) \qquad (6)$$

und

$$V(P_i^+) = \frac{p_i(1-p_i)}{m^+}\,. \qquad (7)$$

Vergleicht man nun ein Stichprobenspektrum $\mathbf{w}^+$ mit einem vorgegebenen Spektrum $\mathbf{w}$, so müßte man nach dieser Überlegung eine größere Abweichung noch als zufällig hinnehmen, als es bei Anwendung des KOLMOGOROFF-SMIRNOV-Tests geschieht. Die Aussage dieses Tests wird aber umso treffender sein, je geringer der Variationskoeffizient $\sqrt{V(R)}/\bar{R}$ der Rechnungsbeträge ist.

Bei großem Stichprobenumfang m^+ sind die W_i^+ annähernd normalverteilt. Dann läßt sich zur Stabilitätsüberwachung der t-Test[2] anwenden, der die Varianz der Komponenten ausdrücklich mitbe-

---

1 Vgl. S. 56.

2 Vgl. L. Neumeyer, a.a.O., S. 94 ff. Bei NEUMEYER werden wegen der besonderen Art der Ermittlung der Verweilzeitverteilung Anteile auf Stabilität getestet, während hier Verhältnisse betrachtet werden.

rücksichtigt und so auch dem besonderen Charakter der Betragsverteilung Rechnung trägt:

Um festzustellen, ob die Kontrollstichprobe aus einer Grundgesamtheit stammt, in der die Betragsverteilung **w** gilt, prüft man, ob die Differenz $w_i^+ - w_i$ für alle Verweilzeiten i nur zufällig von Null verschieden ist. Die Prüfgröße lautet

$$t_i = \frac{w_i^+ - w_i}{\sqrt{V(W_i^+ - w_i)}} \quad , \; i=0,\ldots,n, \tag{8}$$

oder, wegen der Konstanz des fest vorgegebenen Wertes w_i,

$$t_i = \frac{w_i^+ - w_i}{\sqrt{V(W_i^+)}} \quad , \; i=0,\ldots,n, \tag{9}$$

wobei man den Schätzwert für $V(W_i^+)$ als

$$\hat{V}(W_i^+) = \frac{1}{m^+ \cdot r^2} \sum_{j=1}^{m^+} \frac{(b_{ij} - w_i^+ r_j)^2}{m^+ - 1} \tag{10}$$

erhält[1].

Solange nun für keine Verweilzeit i die Prüfgröße t_i den kritischen Wert t_c für $\nu = m^+ - 1$ Freiheitsgrade übersteigt, darf man annehmen, daß die Verweilzeitverteilung stabil geblieben ist. Bei großem Stichprobenumfang ($m^+ > 60$) und zweiseitiger Fragestellung läßt sich der kritische Wert zur Vereinfachung generell als $t_c = 2$ angeben, wenn auf einem Signifikanzniveau von $\alpha = 0{,}05$ geprüft werden soll[2].

---

1 Vgl. S. 49. (Der Auswahlsatz ist als sehr klein angenommen.)
2 Vgl. etwa S. Siegel, a.a.O., S. 248.

Für extrem kleine Komponenten der Betragsverteilung kann man auch bei ausgesprochen großem Stichprobenumfang nicht damit rechnen, daß sich die Verteilung der entsprechenden Stichprobenparameter W_i^+ gut durch die Normalverteilung approximieren läßt. Der t-Test ist zwar recht robust in Bezug auf Abweichungen von der Normalverteilung[1], die Testergebnisse dieser Komponenten sollten jedoch vorsichtig interpretiert werden. Da sehr kleine w_i ohnehin nur einen untergeordneten Einfluß auf das Prognoseergebnis nehmen, kann man eventuell auf eine Neuermittlung des Spektrums verzichten, solange nur solche unbedeutenden Komponenten signifikante Veränderungen aufweisen.

Der t-Test geht von den drei vorgestellten Verfahren der Stabilitätskontrolle am besten auf die tatsächliche Problemstellung ein, und er ist obendrein derjenige mit der größten Trennschärfe[2], d.h. er ist am besten geeignet, Instabilität festzustellen. Daraus darf man jedoch nicht die Forderung ableiten, zur Überwachung der Stabilität grundsätzlich diesen Test zu verwenden. Strenggenommen sind die drei Vorschläge nämlich keine Alternativen.

Das Verfahren von BENISHAY setzt die Ermittlung des Spektrums $\mathbf{w}^+$ aus den Restforderungen voraus. Wird $\mathbf{w}^+$ aus einer Stichprobe von m^+ Rechnungen gewonnen, so sind die einzelnen Komponenten nicht mehr voneinander unabhängig; der Vorzeichen- und der run-Test lassen sich dann nicht anwenden. Der KOLMOGOROFF-SMIRNOV-Test und der t-Test hingegen können gerade nur dann benutzt werden, wenn der Berechnung von $\mathbf{w}^+$ eine Stichprobe zugrunde liegt. Welches der beiden Verfahren in diesem Fall verwendet werden sollte, ist dann auch nicht frei wählbar, sondern richtet sich nach der Größe des Stichprobenumfangs[3].

---

1 Vgl. W. L. Hays, a.a.O., S. 410.

2 Vgl. L. Sachs, a.a.O., S. 127.

3 Dabei wird davon ausgegangen, daß man nicht den weniger günstigen KOLMOGOROFF-SMIRNOV-Test benutzt, wenn der t-Test angewendet werden kann.

Nun kann man fordern, es sollten stets so große Stichproben gezogen werden, daß der für die Aufdeckung fehlender Stabilität optimale t-Test tatsächlich eingesetzt werden kann. Dem steht aber die folgende Überlegung entgegen:

Aufgabe der Stabilitätsüberwachung ist es, durch die Untersuchung regelmäßig entnommener Kontrollstichproben festzustellen, ob das alte Spektrum noch verwendet werden kann oder ob eine Neuberechnung notwendig ist. Da man im voraus nicht weiß, wann sich das Spektrum $\mathbf{w}$ verändern wird, müssen die Stichproben mit möglichst kurzen zeitlichen Abständen gezogen werden[1], damit man nicht zu lange mit der nicht mehr gültigen Verweilzeitverteilung arbeitet. NEUMEYER empfiehlt, die Intervalle zwischen den Entnahmen in der Größenordnung der Basisperiode zu wählen[2]. Das bedeutet, daß z.B. bei wochenweiser Einnahmeprognose in jeder Woche eine Kontrollstichprobe gezogen wird. Folgt man diesem Vorschlag, so verbietet sich - besonders bei kurzen Basisperioden - die regelmäßige Entnahme großer Stichproben schon wegen der dadurch entstehenden hohen Kosten. Dazu tritt noch ein grundsätzlicher Einwand. Die Ermittlung von $\mathbf{w}^+$ aus einer Stichprobe, die so groß ist, daß sie die Anwendung des t-Tests gestattet, stellt praktisch bereits die vollständige Neuberechnung der Verweilzeitverteilung dar und macht den Test somit überflüssig; denn der soll ja gerade bei der Entscheidung darüber helfen, ob eine Neuberechnung erforderlich ist. Wenn in diesem Fall noch ein Stabilitätstest durchgeführt wird, so dient er nicht mehr der Kontrolle im oben genannten Sinne, sondern eher der Analyse des Zahlungsverhaltens der Kunden im Zeitablauf.

Die direkte Stabilitätsüberwachung befindet sich offensichtlich in einem Dilemma. Entweder sie nimmt eine reine Kontrollfunk-

---

1 Vgl. L. Neumeyer, a.a.O., S. 107.

2 Ebenda, S. 108.

tion wahr und gestattet dabei nur die Anwendung weniger geeigneter Tests, oder sie arbeitet mit dem optimalen Test, gibt dann aber ihre eigentliche Funktion auf.

Als Ausweg bietet sich ein Kompromiß an: In den oben angegebenen zeitlichen Intervallen werden kleine Kontrollstichproben gezogen, die wenigstens grobe Abweichungen aufzeigen können und bei Bedarf eine vollständig neue Ermittlung der Betragsverteilung veranlassen. Darüberhinaus werden solche Neuberechnungen in größeren zeitlichen Intervallen ohne Anstoß durch einen Test vorgenommen, damit auch kleinere Verschiebungen der Verteilung erfaßt werden können.

Was ist nun zu tun, wenn das aus einer Kontrollstichprobe resultierende Spektrum w^+ n i c h t signifikant vom bisher verwendeten Spektrum w abweicht? Denkbar wäre es, die alte Betragsverteilung solange weiterzuverwenden, bis eine große Stichprobe gezogen wird und eine wirkliche Neuberechnung erfolgt. Hier soll ein anderer Weg beschritten werden. Wenn auch bei der Schätzung der Betragsverteilung aus einer kleinen Stichprobe der Zufallsfehler zu groß ist, als daß man w^+ anstelle von w für zukünftige Prognosen verwenden sollte, so erscheint es doch nicht zweckmäßig, auf die aus der Kontrollstichprobe gewonnene Information über das Zahlungsverhalten der Kunden vollständig zu verzichten, zumal sie jüngsten Datums ist. Es soll vielmehr gezeigt werden, wie man w^+ bei der weiteren Prognosetätigkeit mitberücksichtigen kann.

Angenommen, in der Periode t sei aus einer umfangreichen Erhebung die Verteilung w_t ermittelt worden. Eine Periode später habe sich aus einer Kontrollstichprobe das von w_t nicht signifikant verschiedene Spektrum w^+_{t+1} ergeben. Da die Annahme der Stabilität nicht abgelehnt wird, liegt es nahe, in t+1 ein neues Spektrum w_{t+1} als gewogenes Mittel aus w_t und w^+_{t+1} zu berechnen,

$$w_{t+1} = A w^{+}_{t+1} + (1-A) w_t \quad \text{mit } 0 < A < 1 , \tag{11}$$

welches dann weiterverwendet wird. Die Mittelwertbildung wird nach

$$w_{t+k} = A w^{+}_{t+k} + (1-A) w_{t+k-1} \tag{12}$$

solange durchgeführt, bis sich ein signifikanter Unterschied zwischen w_{t+k-1} und w^{+}_{t+k} zeigt oder ohnehin eine vollständige Neuberechnung des Spektrums ansteht. Sie entspricht dem BROWNschen exponentiellen Glätten[1] und führt zu einer exponentiellen Gewichtung der aus den Kontrollstichproben ermittelten Spektren bei der Berechnung von w_{t+k}[2]:

$$w_{t+k} = A(1-A)^0 w^{+}_{t+k} + A(1-A)^1 w^{+}_{t+k-1} + A(1-A)^2 w^{+}_{t+k-2} +$$

$$\ldots + A(1-A)^{k-1} w^{+}_{t+1} + (1-A)^k w_t . \tag{13}$$

Die Struktur der Gewichtung wird durch die Glättungskonstante A festgelegt. Wie groß A gewählt werden sollte, läßt sich mit einer verbindlichen Vorschrift nicht angeben[3]. Bei der exponentiellen Fortschreibung eines Mittelwertes hat es sich aber als Erfahrungsregel gezeigt, daß A zwischen 0,1 und 0,3 liegen sollte[4].

---

1 Vgl. R. G. Brown: Smoothing, Forecasting and Prediction of Discrete Time Series, Englewood Cliffs, N. J. 1963, S. 101.

2 Ebenda.

3 Vgl. R. Lewandowski: Prognose- und Informationssysteme und ihre Anwendungen, Band 1, Berlin/New York 1974, S. 37 ff; M. Schröder: Einführung in die kurzfristige Zeitreihenprognose und Vergleich der einzelnen Verfahren, in: P. Mertens (Hrsg.): Prognoserechnung, a.a.O., S. 38ff.

4 Vgl. K. Reif: Bedarfsvorhersage mittels mathematisch-statistischer Verfahren, IBM Form 81 518, S. 9; N. L. Enrick u. W. Schäfer: Quantitative Marktprognose, Heidelberg 1972, S. 26.

Ein zusätzlicher Anhaltspunkt bei der Wahl von A ergibt sich aus dem Vergleich des Umfangs der Kontrollstichproben, der zur Vereinfachung gleichbleibend als m^+ angenommen sei, mit dem Umfang m der Stichprobe, aus der w_t berechnet worden ist. Es sei davon ausgegangen, daß eine Stichprobe von m Rechnungen genügt, um eine befriedigende Abschätzung des Spektrums zu gestatten. E i n e der kleineren Kontrollstichproben reicht dann zwar zu einer eigenständigen Neuberechnung der Verteilung nicht aus, jedoch können $v = m/m^+$ Stichproben des Umfangs m^+ diese Berechnung gestatten, solange Stabilität vorliegt. Da v nicht notwendigerweise ganzzahlig ist, wird noch nach v' aufgerundet.

Will man nun erreichen, daß die v' jüngsten Kontrollstichproben den Hauptanteil, z.B. 95 %, des Gewichts bei der Fortschreibung der Verweilzeitverteilung erhalten, dann ist A so zu wählen, daß

$$\sum_{k=0}^{v'-1} A(1-A)^k = 0{,}95 \qquad (14)$$

erfüllt ist. Aus

$$\sum_{k=0}^{v'-1} A(1-A)^k = A\,\frac{(1-A)^{v'}-1}{(1-A)-1} = 1 - (1-A)^{v'} = 0{,}95$$

ergibt sich der gesuchte Wert für die Glättungskonstante als

$$A = 1 - 0{,}05^{\frac{1}{v'}} . \qquad (15)$$

Nicht nur das Liquidationsspektrum, auch der zur Abgabe von Intervallprognosen benötigte Satz von Varianzen $V(B_i)$, i=0,...,n, kann sich im Zeitablauf verändern. Auf eine eigene Überprüfung der Stabilität von $V(B_i)$, die sich erheblich schwieriger ge-

staltet als die Überwachung des Spektrums[1], sei verzichtet. Entsprechend der Fortschreibung der Verweilzeitverteilung soll aber wenigstens eine Aktualisierung der bisher verwendeten Werte stattfinden:

Die in t aus der großen Stichprobe ermittelten Schätzwerte für die Varianzen seien mit $\hat{V}_t(B_i)$ bezeichnet und die entsprechenden Werte aus der Kontrollstichprobe der Periode t+1 mit $V^+_{t+1}(B_i)$. Die Fortschreibung der Varianzen erfolgt dann nach

$$\hat{V}_{t+1}(B_i) = A\, \hat{V}^+_{t+1}(B_i) + (1-A)\, \hat{V}_t(B_i),\ i=0,\ldots,n, \qquad (16)$$

wobei A nicht unbedingt mit dem Glättungsparameter der Spektren identisch zu sein braucht. Sie wird solange fortgeführt, bis ein neues Spektrum ermittelt wird und dabei auch eine vollständige Neuberechnung der Varianzen $\hat{V}(B_i)$ erfolgt.

5.3 Indirekte Stabilitätsüberwachung

Stabilitätsüberwachung ist nicht nur durch den unmittelbaren Vergleich zweier zu verschiedenen Zeitpunkten ermittelter Verweilzeitverteilungen möglich, sie kann auch indirekt erfolgen, indem man an dem Prognosefehler ansetzt. Bei der Einnahmeprognose mit Hilfe eines Liquidationsspektrums, welches dem Zahlungsverhalten der Kunden nicht mehr entspricht, ergibt sich außer dem zufallsbedingten auch ein systematischer Prognosefehler. Die indirekte Stabilitätsüberwachung geht nun davon aus, daß das verwendete Spektrum gültig ist, solange sich die Abweichungen zwischen prognostizierten und realisierten Einnahmen so verhalten, wie man es bei Stabilität der Verweilzeitverteilung erwarten würde[1], und prüft, ob ein systematischer Fehler vorliegt.

---

1 Wegen der ausgeprägten Schiefe der Verteilungen der B_i kann z.B. auch bei großen Stichproben der übliche Test für Varianzen mit der χ^2-Verteilung nicht durchgeführt werden. Vgl. W. L. Hays, a.a.O., S. 442 f.

2 Vgl. R. Edin: Dynamische Analyse..., a.a.O., S. 60.

Der Prognosefehler der Periode t wurde als Differenz zwischen den tatsächlichen und den für t prognostizierten Einnahmen definiert:

$$f_t = e_t - \hat{e}_t \quad . \tag{1}$$

Wie bei der nachträglichen Beurteilung der Genauigkeit der Prognose stehen auch hier mehrere Werte für $\hat{e}_t$ zur Verfügung, welche in verschiedenen Perioden vor t berechnet worden sind. Während dort aber eine allgemein verbindliche Vorschrift für die Auswahl eines bestimmten dieser Werte nicht gegeben werden konnte, ist jetzt aus der Zielsetzung der Stabilitätsüberwachung zu fordern, daß derjenige Prognosewert $\hat{e}_t$ verwendet werden sollte, zu dessen Berechnung die tatsächlich realisierten Umsätze u_{t-i} herangezogen worden sind[1]. Da die Stabilitätsüberwachung nur Abweichungen erfassen soll, die aus dem veränderten Zahlungsverhalten resultieren, muß nämlich der durch falsch vorhergesagte Umsätze verursachte Teil des Fehlers ausgeschaltet werden.

Eine Prognose der Einnahmen in t, welche ausschließlich bereits realisierte Umsätze berücksichtigt, kann aufgrund der Struktur der Prognosegleichung

$$\hat{e}_t = \sum_{i=0}^{n} u_{t-i} w_i \tag{2}$$

in solchen Fällen, in denen w_0 von Null verschieden ist, nur in t oder später abgegeben werden, also erst dann, wenn e_t schon bekannt ist. Es handelt sich somit um eine Ex-post-Prognose, die eigens zum Zwecke der Stabilitätsüberwachung durchgeführt werden muß, sofern nicht regelmäßig bei der Prognose der nächsten n Periodeneinnahmen auch der Wert der laufenden Periode gleich mitbestimmt wird.

---

1 Vgl. H.J. Schmitt: Planungsbuchhaltung - Ein Beitrag zur Integration von Buchhaltung und Planung, Berlin 1971, S. 111; J. Schwarze, a.a.O., S. 551.

Wie kann man nun anhand des Prognosefehlers verändertes Zahlungsverhalten erkennen? Für die nachträgliche Beurteilung der Prognosegenauigkeit wurden bereits Fehlermaße angegeben, die auf systematische Abweichungen reagieren. Zur Stabilitätsüberwachung sind diese Maße aber nur bedingt geeignet, weil ihre Interpretation viel Erfahrung und Fingerspitzengefühl erfordert. Eine Ausnahme ist lediglich die von LANGEN vorgeschlagene Prüfgröße

$$t = \frac{MA \sqrt{T-1}}{s_{f_t}} \tag{3}$$

in der

$$MA = \frac{1}{T} \sum_{t=1}^{T} f_t \tag{4}$$

der mittlere beobachtete Fehler und

$$s_{f_t} = \sqrt{\frac{1}{T} \sum_{t=1}^{T} (f_t - MA)^2} \tag{5}$$

die Standardabweichung der beobachteten Abweichungen ist, und die mit dem entsprechenden Wert der t-Verteilung verglichen wird. Auf diese Weise wird der Versuch unternommen, sich von der subjektiven Beurteilung unabhängig zu machen.

Es wurde bereits darauf hingewiesen, daß die Bedingungen des t-Tests bei der Berechnung von t nach (3) nicht erfüllt sind. Darauf soll nun näher eingegangen werden. Der t-Test setzt voraus[1], daß die Beobachtungen f_t

- voneinander unabhängig
- aus normalverteilten Grundgesamtheiten
- mit jeweils derselben Varianz $V(F_t)$ stammen.

Daß keine Unabhängigkeit der Fehler vorliegt, kann man aus der Tabelle 11 erkennen.

---

1 Vgl. S. Siegel, a.a.O., S. 19.

Periode	t-3	t-2	t-1	t
Umsatz	u_{t-3}	u_{t-2}	u_{t-1}	u_t
tatsächlich in dieser Periode eingehende Einnahmeteile	$u_{t-3}w_{0,t-3}$ $u_{t-4}w_{1,t-4}$ $u_{t-5}w_{2,t-5}$	$u_{t-2}w_{0,t-2}$ $u_{t-3}w_{1,t-3}$ $u_{t-4}w_{2,t-4}$	$u_{t-1}w_{0,t-1}$ $u_{t-2}w_{1,t-2}$ $u_{t-3}w_{2,t-3}$	$u_t w_{0,t}$ $u_{t-1}w_{1,t-1}$ $u_{t-2}w_{2,t-2}$
Gesamte Einnahme der Periode	e_{t-3}	e_{t-2}	e_{t-1}	e_t

Tabelle 11: Tatsächliche Liquidation der Periodenumsätze

Darin ist $w_{i,t-k}$ der tatsächliche Anteil des Umsatzes der Periode t-k mit der Verweilzeit i. Die durch Striche verbundenen Einnahmeteile sind voneinander unabhängig: Ist ein Teil $u_{t-k}w_{i,t-k}$ größer (kleiner) als der erwartete Wert $u_{t-k}w_i$, so sind die übrigen Einnahmeteile, die aus demselben Umsatz u_{t-k} stammen, tendenziell kleiner (größer) als die entsprechenden erwarteten Beträge; zwischen den Einnahmeteilen eines Streckenzuges besteht also eine negative Korrelation. Dadurch sind auch die tatsächlichen Einnahmen aufeinanderfolgender Perioden und somit auch die Prognosefehler negativ korreliert.

Das Ausmaß der Abhängigkeit zwischen zwei Prognosefehlern richtet sich insbesondere nach ihrem zeitlichen Abstand. Es ist tendenziell am größten bei unmittelbar aufeinanderfolgenden Fehlern - dann wird es durch n gemeinsame Streckenzüge beeinflußt - und am geringsten, wenn zwischen beiden Fehlern n-1 Perioden liegen - dann besteht nur noch ein Einfluß durch die Liquidation eines Umsatzes. Bei noch weiter auseinanderliegenden Fehlern besteht Unabhängigkeit.

Zusammenfassend kann somit gesagt werden: Die Prognosefehler f_t können nicht als unabhängige Beobachtungen aufgefaßt werden. Sie sind vielmehr negativ korreliert. Das Ausmaß der Abhängigkeit nimmt ceteris paribus mit dem zeitlichen Abstand zwischen den Fehlern ab. Die Korrelation hat zur Folge, daß die Prüfgröße tendenziell zu klein wird[1] und eine mögliche Instabilität nicht rechtzeitig erkennt.

Auf die 2. Voraussetzung des t-Tests soll nicht weiter eingegangen werden. Bei der Besprechung der Intervallprognose wurde bereits ausgeführt, daß unter bestimmten Bedingungen - besonders bei großer Anzahl von pro Periode erstellten Rechnungen - mit annähernder Normalverteilung der Fehler gerechnet werden kann. Im übrigen ist der t-Test recht robust gegen Abweichungen von der Normalverteilung.

Dagegen verdient die 3. Annahme wieder besondere Aufmerksamkeit. Der LANGENsche Stabilitätstest unterstellt gleiche Varianz für die Fehler verschiedener Perioden. Die ist jedoch selbst bei Stabilität des Liquidationsspektrums nicht gegeben. Die Varianz des Prognosefehlers der Periode t ergibt sich vielmehr als

$$V(F_t) \doteq \sum_{i=0}^{n} m_{t-i} \, V(B_i) \; . \qquad (6)$$

Die Varianzen $V(B_i)$ sind von der Streuung und der mittleren Höhe der Rechnungsbeträge und dem Zahlungsverhalten der Kunden abhängig[2]. Selbst wenn sie als konstant angenommen werden können, hängt $V(F_t)$, genau wie $\hat{e}_t$, von der Anzahl der in den Perioden t-n bis t erstellten Rechnungen ab. Nur bei ständig gleichbleibender Anzahl m_{t-i} ist dann auch $V(F_t)$ über alle Perioden hinweg konstant. Dieser Fall ist aber sehr unrealistisch. Die Aussagekraft

---

1 Vgl. R. H. Woodward u. P. L. Goldsmith: Cumulative Sum Techniques, I. C. I. Monograph No. 3, Edinburgh 1964, S. 18.

2 Vgl. S. 96.

dieses Tests ist somit begrenzt. Sie ist umso größer, je länger das Liquidationsspektrum ist und je weniger die Anzahl der pro Periode erstellten Rechnungen - etwas ungenauer: je weniger der Umsatz pro Periode - schwankt.

WRIGHT versucht, der Heteroskedastizität der Prognosefehler dadurch Rechnung zu tragen, daß er nicht die absoluten Fehler f_t, sondern die prozentualen Fehler

$$d_t = \frac{f_t}{\hat{e}_t} \cdot 100\% \qquad (7)$$

betrachtet[1]. Er schlägt vor, die relativen Abweichungen d_t auf einer Qualitätskontrollkarte abzutragen und die Varianz s_d^2 dieser Abweichungen zu ermitteln. Eine ungünstige Entwicklung[2] der Verweilzeitverteilung wird auf einem Signifikanzniveau von $\alpha = 0{,}05$ dann angenommen, wenn ein Wert von d_t auftritt, der kleiner ist als $-1{,}65\ s_d$, ferner wenn fünf aufeinanderfolgende Abweichungen negativ sind oder wenn drei aufeinanderfolgende Differenzen $(d_{t+1}-d_t)$ ein negatives Vorzeichen tragen[3].

Die einseitige Fragestellung erscheint nicht angebracht. Wenn sich das Spektrum verändert hat, so ist es - auch wenn die Veränderung günstig ausfällt - zweckmäßig, eine neue Verweilzeitverteilung zu ermitteln, damit unverzerrte Prognosen abgegeben werden können. Eine Unterschätzung der Einnahmen kann nämlich zur Haltung unnötig hoher Liquiditätsreserven und damit zu vermeidbaren Kosten führen. Die Erweiterung des WRIGHTschen Ansatzes zu einem zweiseitigen Test ist aber ohne Schwierigkeiten möglich. Die Stabilitätshypothese wird dann auf einem Signi-

---

1 Vgl. F. K. Wright: Managerial Controls of Accounts Receivable: A Comment, Journal of Accounting Research 4(1966), S. 128. WRIGHT betrachtet nicht die Prognose der Einnahmen, sondern die der Restforderungen. Der Stabilitätstest läßt sich aber unmittelbar übertragen.

2 WRIGHT prüft nur, ob das Spektrum ungünstiger geworden ist.

3 Vgl. F. K. Wright, a.a.O., S. 128.

fikanzniveau von $\alpha = 0{,}05$ verworfen, wenn die Intervallgrenzen $\pm 1{,}96\ s_d$ überschritten werden, wenn sechs aufeinanderfolgende Abweichungen dasselbe Vorzeichen tragen[1] oder wenn für fünf aufeinanderfolgende Fehler $d_{t-4} < d_{t-3} < d_{t-2} < d_{t-1} < d_t$ bzw. $d_{t-4} > d_{t-3} > d_{t-2} > d_{t-1} > d_t$ gilt[2].

Auch dieser Test geht von der nicht zutreffenden Annahme unabhängiger Prognosefehler aus. Ferner wird der durch die Relativierung der Fehler gewünschte Effekt der Homogenisierung der Fehlervarianzen nicht erreicht. Es gilt nämlich

$$V(D_t) = V\left(\frac{F_t}{\hat{e}_t}\right) = \frac{1}{\hat{e}_t^2}\, V(F_t), \tag{8}$$

und da sowohl $V(F_t)$ wegen

$$V(F_t) \doteq \sum_{i=0}^{n} m_{t-i}\, V(B_i) \tag{9}$$

als auch $\hat{e}_t$ wegen

$$\hat{e}_t = \sum_{i=0}^{n} m_{t-i} w_i \tag{10}$$

eine lineare Funktion von m_{t-i} $(i=0,\ldots,n)$ ist, sinkt $V(D_t)$ mit zunehmendem $\hat{e}_t$. Man kann sich deshalb die Relativierung der Prognosefehler ersparen und unmittelbar die Ausgangswerte heranziehen. Soweit es den t-Test betrifft, besteht somit kein prinzipieller Unterschied zwischen dem LANGENschen Vorschlag und dem von WRIGHT. Die Stabilitätsüberwachung nach Art der Kontrollkartentechnik gibt aber über den t-Test hinaus einen Einblick in die Entwicklung der Fehler im Zeitablauf und erleichtert damit die Interpretation, und da die Kontrollkartentechnik ohnehin eher eine Entscheidungshilfe als ein Instrument zur Abschätzung ge-

---

1 Vgl. E. L. Grant: Statistical Quality Control, 3rd edition, New York 1964, S. 95; D. A. Simmons: Practical Quality Control, Reading 1970, S. 136.

2 Vgl. P. S. Olmstaed: Distribution of Sample Arrangements for Runs up and down, The Annals of Mathematical Statistics 17 (1946), S. 27.

nauer Wahrscheinlichkeiten sein will[1], ist der Ansatz von WRIGHT vorzuziehen.

Die WRIGHTsche Stabilitätsüberwachung ist, wie auch die nach LANGEN, auf die Annahme der Homoskedastizität der betrachteten Fehler angewiesen, weil die Varianz aus den beobachteten Fehlern geschätzt werden muß. Sofern die Anzahl der pro Periode erstellten Rechnungen im Zeitablauf um einen mittleren Wert oszilliert, ist diese Annahme wenigstens annähernd erfüllt. Bei zeitlich konstanten Varianzen $V(B_i)$ (i=0,...,n) übt die Beziehung

$$V(F_t) \doteq \sum_{i=0}^{n} m_{t-i} \, V(B_i) \qquad (11)$$

nämlich eine ähnlich glättende Wirkung auf die Zeitreihe der Varianzen $V(F_t)$ aus wie die Bildung gleitender Mittelwerte. Im günstigen Fall werden die $V(F_t)$ nur geringfügig um einen mittleren Wert streuen, der nach

$$\hat{V}(F_t) = \frac{1}{T-1} \sum_{t=1}^{T} \left(f_t - \frac{1}{T} \sum_{t=1}^{T} f_t\right)^2 \qquad (12)$$

geschätzt wird.

Sobald jedoch eine ständige deutliche Umsatzausweitung vorliegt oder ausgeprägte Saisonschwankungen auftreten, deren Periodenlänge die Länge der Verweilzeitverteilung übertrifft[2], wird sich auch die Varianz $V(F_t)$ auf entsprechende Weise im Zeitablauf verändern. Es ist dann nicht zweckmäßig, mit einer aus den beobachteten Fehlern f_t nach (12) geschätzten mittleren Varianz zu arbeiten.

---

1 Vgl. E. L. Grant, a.a.O., S. 221; Western Electric: Statistische Qualitätskontrolle - Das Handbuch der Western Electric Company für den Praktiker, Stuttgart 1969, S. 1.

2 Wenn am Ende des Spektrums viele kleine Komponenten auftreten, ist weniger seine gesamte Länge entscheidend als die Anzahl derjenigen Komponenten, die deutlich von Null verschieden sind. Für sehr kleine w_i ist nämlich die zugehörige Varianz $V(B_i)$ ebenfalls gering und der Beitrag $m_{t-i}V(B_i)$ zur Varianz $V(F_t)$ nur unbedeutend (vgl. S. 96).

Probleme können bei der Schätzung nach (12) auch auftreten, wenn die Basisperiode sehr lang gewählt ist, etwa wenn Monatsdaten betrachtet werden. Um eine genügend große Anzahl von Beobachtungen f_t zur Berechnung der Varianz zu erhalten, muß man dann auch weit zurückliegende Fehler heranziehen, die u. U. gar nicht mehr repräsentativ für das Zahlungsverhalten der Kunden sind.

Die Annahme homogener Fehlervarianzen soll nun aufgehoben werden. Liegen Schätzwerte für die Varianzen $V(B_i)$ vor - das ist dann der Fall, wenn das Spektrum nach der Verhältnisschätzung ermittelt wurde - so läßt sich für jeden Fehler F_t die individuelle Varianz $V(F_t)$ als

$$\hat{V}(F_t) = \sum_{i=0}^{n} m_{t-i} \hat{V}(B_i) \qquad (13)$$

angeben, und auf der Kontrollkarte, auf der die Fehler f_t abgetragen werden, erhält man individuell berechnete, gestufte Kontrollgrenzen, deren Abstand von der Linie $f_t=0$ jeweils $\pm t \sqrt{\hat{V}(F_t)}$ beträgt[1]. Bei stark saisonal schwankenden Umsätzen können sich z.B. Kontrollgrenzen nach Art der Abbildung 14 ergeben.

Aus einer Kontrollkarte dieser Art läßt sich zwar feststellen, ob ein Fehler außerhalb des entsprechenden Intervalls liegt und welches Vorzeichen er trägt, ein direkter Betragsvergleich der f_t ist jedoch wenig sinnvoll. Da gestufte Kontrollgrenzen überdies nicht so übersichtlich sind wie parallel zu $f_t=0$ verlaufende Schranken[2], soll noch eine Standardisierung der Fehler nach

$$f'_t = \frac{f_t}{\sqrt{\hat{V}(F_t)}} \qquad (14)$$

erfolgen.

---

1 t findet man in der t-Verteilung mit m-n-3 Freiheitsgraden, wobei m der Umfang der Stichprobe ist, aus der **w** und die Varianzen $V(B_i)$ ermittelt worden sind, und n die längste Verweilzeit angibt.[1] Man kann aber auch bei $\alpha = 0{,}05$ zur Vereinfachung wieder $t=2$ setzen.

2 Vgl. Western Electric, a.a.O., S. 14.

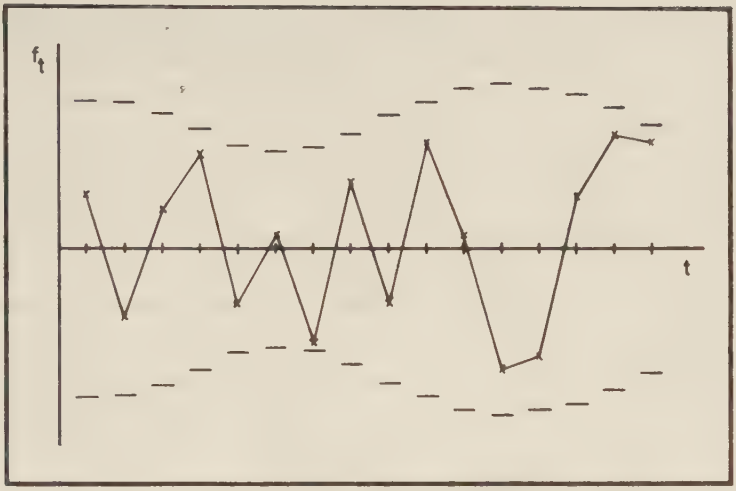

Abb. 14: Beispiel für gestufte Kontrollgrenzen

Die Kontrollgrenzen verlaufen dann im Abstand $t = \pm 2$ von der Linie $f'_t=0$ und die standardisierten Abweichungen f'_t können untereinander auch betragsmäßig verglichen werden. Die Abbildung 15 auf Seite 144 ist durch Standardisierung aus Abbildung 14 entstanden. Man sieht, daß am Ende eine ansteigende Tendenz vorliegt, die aus Abbildung 14 nicht ersichtlich ist.

Der zur Berechnung der Kontrollgrenzen bzw. zur Standardisierung verwendete Wert $\hat{V}(F_t)$ ist ein Schätzwert für die in der Periode t gültige Fehlervarianz und berechnet sich als

$$\hat{V}(F_t) = \sum_{i=0}^{n} m_{t-i} \hat{V}(B_i) \quad . \tag{15}$$

Bisher wurde stillschweigend unterstellt, daß $\hat{V}(B_i)$ - ein aus einer früher gezogenen Stichprobe gewonnener Schätzwert für die Varianz $V(B_i)$ - auch zur Abschätzung der Streuung der in t mit der Verweilzeit i eingehenden Beträge $b_{ij,t-i}$ verwendet werden kann, d.h. daß $V(B_i)$ zeitlich stabil ist. Bei unverändertem Zahlungsverhalten der Kunden mag dieses Annahme realistisch sein;

sobald sich das Liquidationsspektrum verändert, wird sie jedoch von zufälligen Ausnahmen abgesehen nicht mehr zutreffen, weil die Varianzen $V(B_i)$ (i=0,...,n) unter anderem auch von der gültigen Betragsverteilung **w** abhängen. Der mit Hilfe von falsch geschätzten Varianzen $V(B_i)$ berechnete Schätzwert $\hat{V}(F_t)$ für die individuelle Fehlervarianz weicht dann systematisch vom wahren Wert $V(F_t)$ ab.

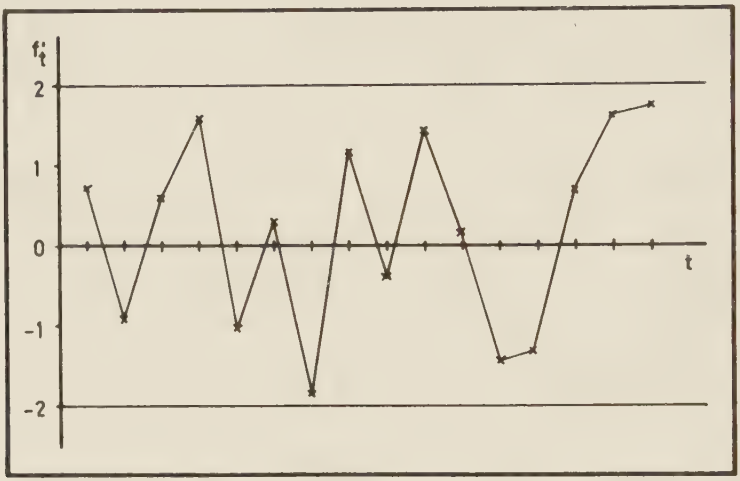

Abb. 15: Kontrollgrenzen der standardisierten Prognosefehler

Ist der verwendete Wert $\hat{V}(F_t)$ kleiner als $V(F_t)$, so wird das Prognoseintervall für die Einnahmen zu günstig eingeschätzt. Beim *t*-Test wird dann der Annahmebereich für die Hypothese H_0 : "**w** ist stabil" zu klein und der für H_1 : "**w** hat sich verändert" zu groß gewählt (vgl. Abbildung 16).

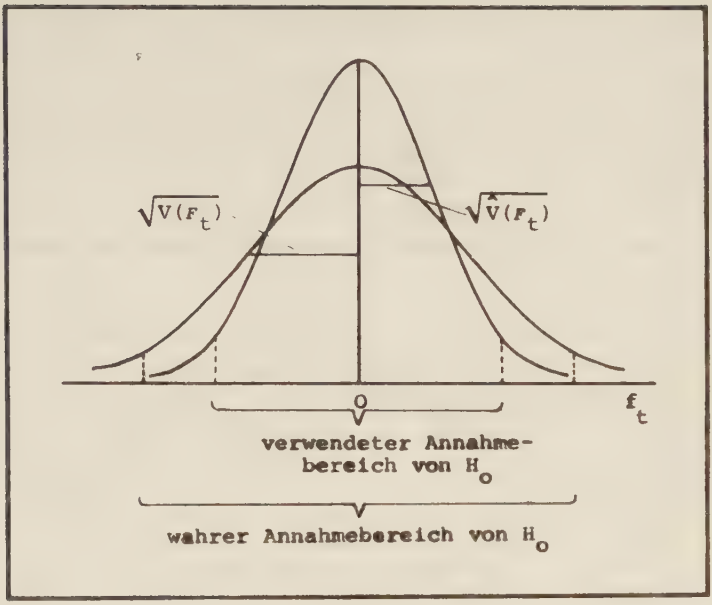

Abb. 16: Veränderung des Annahmebereichs von H_0 bei zu klein geschätzter Varianz $V(F_t)$

Je größer nun die Differenz $V(F_t) - \hat{V}(F_t)$ ist, umso eher wird man H_0 zu Unrecht verwerfen und eine Neuermittlung des Spektrums vornehmen. Das ist durchaus kein Nachteil, im Gegenteil: Da mit der Neuermittlung von **W** auch die Varianzen $V(B_i)$ neu bestimmt werden, bietet sich auf diese Weise die Möglichkeit, auch die Veränderung der $V(B_i)$ implizit im Test zu berücksichtigen[1].

Das ist deshalb wichtig, weil diese Varianzen zur Intervallprognose und zur weiteren Stabilitätsüberwachung benötigt werden. Der Test ist somit in der Lage, eine Verschlechterung der Prognose,

1 Vgl. A. H. Schaafsma u. F. G. Willemze: Moderne Qualitätskontrolle - Statistische und organisatorische Grundlagen der Qualitätsgestaltung, 4. Auflage, Eindhoven 1961, S. 146.

die durch Instabilität des Spektrums bei Gleichbleiben oder Vergrößerung des Prognoseintervalls entsteht, zu erkennen.

Wenn die Fehlervarianz $V(F_t)$ überschätzt wird, ist das Intervall kleiner als angenommen, und der Annahmebereich von H_o wird zu groß gewählt (vgl. Abbildung 17).

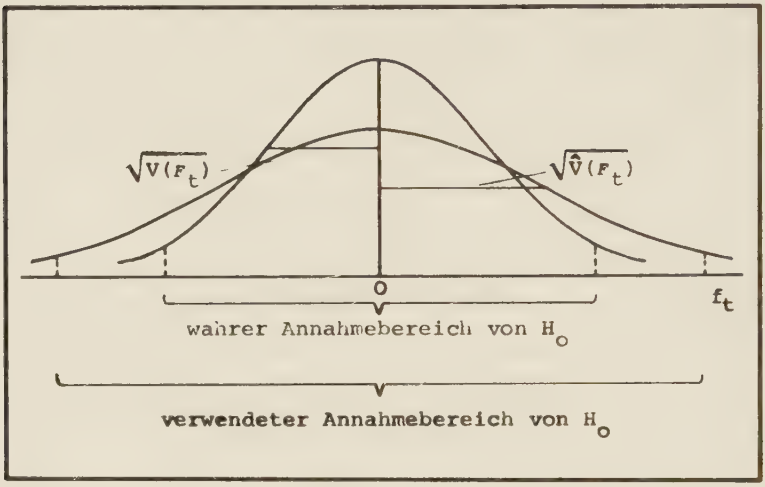

Abb. 17: Veränderung des Annahmebereichs von H_o bei zu groß geschätzter Varianz $V(F_t)$

Die Prognosegenauigkeit hat sich dann in bezug auf die Fehlervarianz erhöht, und diese Verbesserung kann eine gleichzeitig durch Veränderung des Spektrums eingetretene Verzerrung überdecken. Die bei zu klein geschätzter Varianz positiv wirkende Eigenschaft des Tests, sowohl auf Verschiebung des Spektrums als auch auf Veränderung der Varianz $V(F_t)$ zu reagieren, ist hier nachteilig, da nicht zwischen diesen beiden Einflußgrößen unterschieden werden kann.

Je kleiner $V(F_t)$ tatsächlich im Vergleich zu $\hat{V}(F_t)$ ist, umso extremer muß die Verzerrung sein, damit f_t die verwendeten Kontrollgrenzen überschreitet und umso seltener wird eine Veränderung des Spektrums auf diese Weise aufgedeckt werden können. Dafür wächst jedoch mit kleiner werdendem $V(F_t)$ die Chance, daß sich die Instabilität durch längere Folgen von positiven oder negativen Abweichungen äußert. Abbildung 18 soll diesen Zusammenhang verdeutlichen.

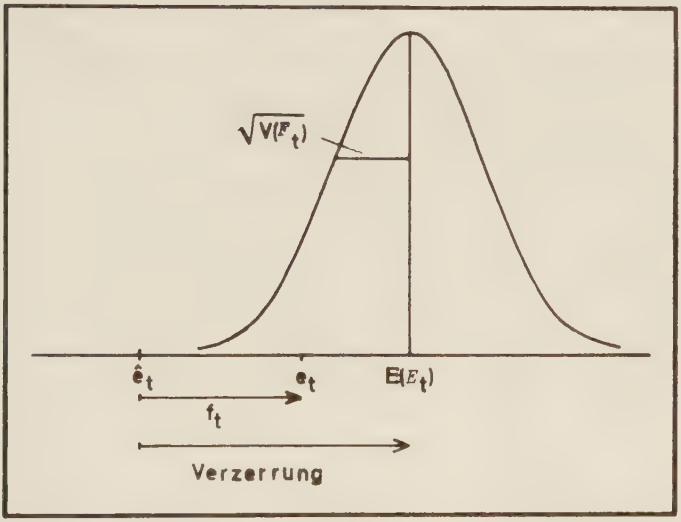

Abb. 18: Zusammenhang von Prognosefehler, Verzerrung und Fehlervarianz

Die Verzerrung der Prognose ergibt sich als Differenz aus dem Erwartungswert der Einnahmen in t und der abgegebenen Prognose $\hat{e}_t$. Je größer nun die Verzerrung ist und je kleiner die Varianz $V(F_t)$, umso seltener wird f_t zufallsbedingt ein "falsches" Vorzeichen tragen, und umso eher wird der Vorzeichentest ansprechen, der somit eine Ergänzung zur Betrachtung der Kontrollgrenzen bedeutet.

Auf den Test der Länge auf- oder absteigender Fehlerfolgen soll nicht näher eingegangen werden. Da er auf standardisierte Fehler angewiesen ist, wird er ebenfalls durch Schätzfehler in den Varianzen $\hat{V}(B_i)$ beeinflußt. Die Auswirkung ist aber nur schwer zu erfassen, und er soll deshalb nur noch als ergänzende Angabe betrachtet werden, die zumindest Tendenzen anzeigt.

6 Einsatz der multiplen Regressionsanalyse unter den Annahmen der Prognose mittels Verweilzeitverteilungen

6.1 Vorbemerkungen

Es wurde bereits gezeigt[1], daß man die in der Periode t eingehenden Umsatzeinnahmen e_t durch die Gleichung

$$e_t = W_0 u_t + W_1 u_{t-1} + \ldots + W_n u_{t-n} + f_t \qquad (1)$$

beschreiben kann, in der

e_t die Ausprägung einer Zufallsvariablen E_t ist,

$W_0 \ldots W_n$ als Betragsverteilung der Grundgesamtheit das mittlere Zahlungsverhalten der Kunden angibt,

$u_t \ldots u_{t-n}$ ein Satz vorgegebener Periodenumsätze ist und

f_t die Ausprägung des Prognosefehlers F_t bei Verwendung der Verweilzeitverteilung $(W_0 \ldots W_n)$ ist, der bei genügend großer Anzahl von Rechnungserstellungen pro Periode annähernd normalverteilt ist und praktisch den Erwartungswert Null hat:

$$E(F_t) \triangleq 0 \; . \qquad (2)$$

Die formale Struktur der Gleichung (1) legt die Frage nahe, ob es nicht möglich ist, mit Hilfe der multiplen Regressionsanalyse (im folgenden kürzer mit MRA bezeichnet) die Parameter $W_0, \ldots, W_n$ zu schätzen und Prognosen für die Umsatzeinnahmen abzugeben[2]. Dieser Frage nachzugehen, erscheint schon deshalb an-

---

1 Vgl. Abschnitt 4.2.1.

2 Vgl. einen ähnlichen Vorschlag bei R. Edin: Dynamische Analyse..., a.a.O., S. 174 ff. EDIN setzt nicht an der Gleichung (1) an, sondern formuliert einen Ansatz ohne Verwendung der Verweilzeitverteilung.

gebracht, weil die Regressionsanalyse im Gegensatz zur Verweilzeitverteilung in der Praxis weitgehend bekannt ist und bereits teilweise angewendet wird[1]. Die Anwendungsgebiete werden bisher zwar schwerpunktmäßig im Absatz- und Fertigungsbereich gesehen[2], wird aber in einem Unternehmen die MRA bereits angewendet, sollte es keine Schwierigkeiten bereiten, sie auch zur Einnahmeprognose heranzuziehen, sofern sie dazu geeignet erscheint.

Eine notwendige technische Voraussetzung für die Anwendbarkeit der MRA ist, daß außer der Umsatzreihe auch die Zeitreihe der realisierten Einnahmen aus Umsätzen vorliegt. Die nachträgliche separate Erfassung der Periodeneinnahmen aus Umsätzen ist nicht selbstverständlich[3] und wird bei der Ermittlung strenggenommen nicht vorausgesetzt. Eine sinnvolle Anwendung der Verweilzeitverteilung schließt jedoch auch die Überwachung des Prognosefehlers ein und erfordert dazu die Kenntnis der realisierten Einnahmen e_t. Somit werden im Grunde keine zusätzlichen Anforderungen an die Organisation der Daten gestellt.

Im weiteren Verlauf sollen die Annahmen über das Zahlungsverhalten der Kunden, welche auch bei der Behandlung der Verweilzeitverteilung unterstellt werden, beibehalten werden. Von besonderer Bedeutung sind dabei die Auswirkungen der Annahmen auf den

---

1 Vgl. J. C. Chambers, S. K. Mullick u. D. D. Smith: An Executive Guide to Forecasting, New York 1974, S. 203; H. Lange, a.a.O., S. 81 und S. 262 ff; H. J. Rogge, a.a.O., S. 182 f; W. Schütz: Methoden der mittel- und langfristigen Prognose - Eine Einführung, München 1975, S. 178; J. Turner: Forecasting Practices in British Industry, London 1974, S. 85 ff.

2 Vgl. Fußnote 1, ferner H. Graf: Die kurzfristige Absatzprognose als Bestandteil der Planung in der Konsumgüterindustrie, Diss. Zürich 1968, S. 154 ff; J. H. Heward u. P. M. Steele: Business Control trough Multiple Regression Analysis - A Technique for the Numerate Manager, Haverhill 1973, S. 23 ff; W. R. Recksiegel: Die Anwendung der Regressions- und Korrelationsanalyse in der Kostenrechnung, Diss. Münster 1972.

3 Vgl. O. Grünewälder: Mittel der Liquiditätskontrolle in industriellen Unternehmungen, Diss. Mannheim 1965, S. 69 ff.

Prognosefehler F_t' und ihre Folgen für die Kleinste-Quadrate-Schätzung der Parameter. Es bleibt anzumerken, daß eine geschlossene Darstellung der MRA nicht gegeben wird. Dazu sei auf die entsprechende Literatur hingewiesen[1]. Hier sollen nur diejenigen Teilaspekte herausgegriffen werden, die aus der speziellen Anwendung heraus Besonderheiten aufweisen.

6.2 Schätzung der Regressionsparameter

Die Gleichung

$$e_t = W_0 u_t + W_1 u_{t-1} + \ldots + W_n u_{t-n} + f_t \qquad (1)$$

kann nicht unmittelbar der Berechnung der Schätzwerte zugrunde gelegt werden. Die Anzahl der in ihr enthaltenen Parameter W_i wird durch n, die längste überhaupt vorkommende Verweilzeit, festgelegt. Nun ist n in aller Regel unbekannt, und damit weiß man auch nicht, wie viele erklärende Variablen tatsächlich in den Ansatz aufgenommen werden sollen.

Selbst bei bekanntem n wäre es darüberhinaus auch gar nicht erstrebenswert, alle Parameter zu berechnen. Bei der Behandlung der Verweilzeitverteilung wurde darauf hingewiesen, daß es zweckmäßig ist, Liquidationsspektren zu stutzen, weil in der Praxis häufig am Ende eines Spektrums viele Komponenten mit kleinen Werten auftreten, die bei der Prognose große Rechenarbeit verursachen, ohne einen wesentlichen Einfluß auf das Ergebnis zu nehmen. Wenn die Ermittlung der Komponenten mit der MRA erfolgen soll, erscheint es angebracht, diese kleinen Parameter gar nicht erst zu berechnen, da der Rechenaufwand mit zunehmen-

---

1 Vgl. etwa A. S. Goldberger: Econometric Theory, New York/London/Sydney 1965; J. Johnston: Econometric Methods, 2nd edition, Tokyo 1972; J. Kmenta, a.a.O.; H. Schneeweiß: Ökonometrie, Würzburg/Wien 1971.

der Anzahl der erklärenden Variablen[1] sehr schnell ansteigt.

In der Gleichung

$$e_t = W_0 u_t + W_1 u_{t-1} + \ldots + W_k u_{t-k} + W_{k+1} u_{t-k-1} + \ldots + W_n u_{t-n} + f_t \qquad (2)$$

sei der Beitrag von $\sum_{i=k+1}^{n} W_i u_{t-i}$ zur Erklärung der Einnahme e_t im Vergleich zu dem von $\sum_{i=0}^{k} W_i u_{t-i}$ nur sehr gering (z. B. über alle Perioden t im Mittel 5 % der Einnahmen). Wie groß k tatsächlich gewählt werden soll, ist der subjektiven Entscheidung überlassen. Es wird zunächst unterstellt, daß ein Wert für k irgendwie vorgegeben ist; auf die Ermittlung wird später eingegangen.

Setzt man nun

$$c_t = \sum_{i=k+1}^{n} W_i u_{t-i} \,, \qquad (3)$$

so läßt sich (2) auch als

$$e_t = c_t + W_0 u_t + W_1 u_{t-1} + \ldots + W_k u_{t-k} + f_t \qquad (4)$$

schreiben. Dabei steht c_t für diejenigen Teile der Einnahmen in t, deren Verursachung länger als k Perioden zurückliegt. Darunter sind insbesondere auch solche Einnahmeteile zu finden, die

---

1 Im eigentlichen Sinne liegt nur eine einzige Variable zur Erklärung der Einnahmen vor, nämlich die Variable "Periodenumsatz". Um formale Übereinstimmung mit dem multiplen Regressionsansatz zu erzielen, soll im folgenden jedoch eine andere Interpretation der Gleichung (1) gewählt werden. Die Einnahmen einer Periode sollen danach durch die Variablen "Umsatz derselben Periode", Umsatz der Vorperiode", usw. erklärt werden. An der rein technischen Durchführung ändert sich dadurch nichts, während die Formulierung erleichtert wird.

aufgrund von Mahnungen hereinfließen oder auf dem Gerichtswege eingezogen worden sind und die hinsichtlich der Verweilzeit schon wegen ihrer geringen Anzahl keine ausgeprägte Regelmäßigkeit aufweisen werden. Bei nicht allzu stürmischer Umsatzausweitung (oder -verringerung) macht man darum sicherlich keinen großen Fehler, wenn man davon ausgeht, daß die Einnahmekomponente c_t die Ausprägung einer Zufallsvariablen ist, welche - von b e s t i m m t e n , vor der Periode t-k getätigten Umsätzen weitgehend unabhängig - im Zeitablauf um einen mittleren Wert W_c zufällig streut, d. h.

$$c_t = W_c + f_t^+ \qquad \text{mit } E(F_t^+) = 0, \tag{5}$$

und das umso weniger, als c_t, gemessen an e_t, ohnehin sehr klein ist. Durch Einsetzen in (4) gelangt man hiermit zu

$$e_t = W_c + W_0 u_t + \ldots + W_k u_{t-k} + f_t + f_t^+ \tag{6}$$

und mit

$$f_t' = f_t + f_t^+ \tag{7}$$

zu

$$e_t = W_c + W_0 u_t + \ldots + W_k u_{t-k} + f_t' \,. \tag{8}$$

Gewöhnlich wird sich f_t' nicht allzu stark von f_t unterscheiden. Da c_t definitionsgemäß sehr klein sein soll, gilt das nämlich erst recht für f_t^+. Dazu kommt noch ein Effekt, der sich aus der formalen Definition von c_t gemäß (3) ergibt:

Der Ausdruck $\sum_{i=k+1}^{n} W_i u_{t-i}$ enthält im Gegensatz zu $\sum_{i=0}^{k} W_i u_{t-i}$ zwar nur kleine, dafür aber sehr viele[1] Komponenten W_i. Die

---

1 Die Notwendigkeit der Verkürzung entsteht gerade durch die große Anzahl dieser unbedeutenden Komponenten.

Summation übt nun, ähnlich der Berechnung gleitender Mittelwerte, eine glättende Wirkung aus, die c_t kaum von W_c abweichen lassen und somit nur kleine f_t^+ erzeugen wird. Wegen der großen Anzahl der Summanden kann man ferner die Verteilung von F_t^+ durch die Normalverteilung approximieren.

Damit ist schließlich der Fehler F_t' wie zuvor F_t bei großer Anzahl von Rechnungserstellungen pro Periode annähernd normalverteilt, und sein Erwartungswert ist bis auf eine unbedeutende Verzerrung gleich Null. Die Gleichung (8) soll der MRA zugrunde gelegt werden.

Die Schätzwerte $\hat{W}_c$, $\hat{W}_0$, ..., $\hat{W}_k$ werden nun so ermittelt, daß in dem Gleichungssystem

$$\begin{aligned}
e_{t-T} &= \hat{W}_c + \hat{W}_0 u_{t-T} + \hat{W}_1 u_{t-T-1} + \ldots + \hat{W}_k u_{t-T-k} + r_{t-T} \\
e_{t-T+1} &= \hat{W}_c + \hat{W}_0 u_{t-T+1} + \hat{W}_1 u_{t-T} + \ldots + \hat{W}_k u_{t-T-k+1} + r_{t-T+1} \\
&\vdots \\
e_t &= \hat{W}_c + \hat{W}_0 u_t + \hat{W}_1 u_{t-1} + \ldots + \hat{W}_k u_{t-k} + r_t
\end{aligned} \tag{9}$$

die Quadratsumme der Residuen $r_j (j= t-T,\ldots,t)$, welche sich als

$$r_j = e_j - \hat{e}_j = e_j - (\hat{W}_c + \sum_{i=0}^{k} \hat{W}_i u_{j-i}) \tag{10}$$

ergeben, ihr Minimum annimmt:

$$\sum_{j=t-T}^{t} r_j^2 \stackrel{!}{=} \min.$$

Dazu definiert man zweckmäßigerweise

$$\mathbf{e} = \begin{pmatrix} e_{t-T} \\ e_{t-T+1} \\ . \\ . \\ . \\ e_t \end{pmatrix} \qquad \mathbf{U} = \begin{pmatrix} 1 & u_{t-T} & u_{t-T-1} & \cdots & u_{t-T-k} \\ 1 & u_{t-T+1} & u_{t-T} & \cdots & u_{t-T-k+1} \\ . & . & . & & . \\ . & . & . & & . \\ . & . & . & & . \\ 1 & u_t & u_{t-1} & \cdots & u_{t-k} \end{pmatrix}$$

$$\hat{\mathbf{w}} = \begin{pmatrix} \hat{w}_c \\ \hat{w}_0 \\ . \\ . \\ . \\ \hat{w}_k \end{pmatrix} \qquad \mathbf{r} = \begin{pmatrix} r_{t-T} \\ r_{t-T+1} \\ . \\ . \\ . \\ r_t \end{pmatrix} .$$

Das Gleichungssystem (9) vereinfacht sich dann zu

$$\mathbf{e} = \mathbf{U}\hat{\mathbf{w}} + \mathbf{r} \quad , \tag{11}$$

und aus

$$\sum_{j=t-T}^{t} r_j^2 = \mathbf{r}'\mathbf{r} \stackrel{!}{=} \min.$$

erhält man die gesuchten Schätzwerte als[1]

$$\hat{\mathbf{w}} = (\mathbf{U}'\mathbf{U})^{-1}\mathbf{U}'\mathbf{e} \quad . \tag{12}$$

Gleichung (12) liefert beste, unverzerrte Schätzwerte für die Parameter der Grundgesamtheit, wenn die Annahmen des klassischen Regressionssatzes erfüllt sind, die hier folgendermaßen lauten[2]:

---

1 Vgl. etwa J. Johnston, a.a.O., S. 123 f.

2 Ebenda, S. 122.

1. Der Erwartungswert des Prognosefehlers F_t' ist in jeder Periode Null.
2. Die Varianz des Prognosefehlers ist über die Perioden hinweg konstant.
3. Prognosefehler verschiedener Perioden sind voneinander unabhängig.
4. Die Matrix $\mathbf{U}$ enthält einen Satz fester Werte.
5. Der Rang von $\mathbf{U}$ ist k+2 und die Anzahl der Beobachtungen ist größer als k+2, d.h. T+2 > k+2.

Das trifft jedoch nicht zu. Die Prämissen eins bis drei sind nicht mit den Aussagen vereinbar, welche aus den Annahmen über das Zahlungsverhalten der Kunden abgeleitet worden sind. Die Auswirkungen dieser Diskrepanz werden im folgenden Abschnitt behandelt. Die vierte Prämisse wirft keine Probleme auf, Schwierigkeiten können dagegen im Rahmen der fünften auftreten, denen ein weiterer Abschnitt gewidmet ist.

6.3 Auswirkung der Annahmen über das Zahlungsverhalten auf die Schätzung

Der Erwartungswert von $\hat{\mathbf{w}}$ beträgt

$$E(\hat{\mathbf{w}}) = E\left[(\mathbf{U}'\mathbf{U})^{-1}\mathbf{U}'\mathbf{e}\right] \qquad (1)$$

oder, da $\mathbf{U}$ eine Matrix fester Werte ist,

$$E(\hat{\mathbf{w}}) = (\mathbf{U}'\mathbf{U})^{-1}\mathbf{U}'E(\mathbf{e}) \; . \qquad (2)$$

Bei großer Anzahl von Rechnungserstellungen pro Periode ist

$$E(E_t) \doteq W_c + W_0 u_t + \ldots + W_k u_{t-k} \qquad (3)$$

und damit

$$E(\mathbf{e}) \doteq \mathbf{U}\mathbf{w} \tag{4}$$

mit

$$\mathbf{w} = \begin{pmatrix} W_C \\ W_O \\ \cdot \\ \cdot \\ \cdot \\ W_k \end{pmatrix} . \tag{5}$$

Einsetzen in (2) führt zu

$$E(\hat{\mathbf{w}}) \doteq (\mathbf{U}'\mathbf{U})^{-1}\mathbf{U}'\mathbf{U}\mathbf{w} = \mathbf{w} . \tag{6}$$

Die Kleinste-Quadrate-Regression liefert somit bei großer Anzahl von Fällen pro Periode praktisch unverzerrte Schätzwerte für die Betragsverteilung.

Für eine Intervallschätzung der Komponenten W_i benötigt man Schätzwerte für die Varianzen $V(\hat{W}_i)$, die im klassischen Modellansatz der MRA folgendermaßen berechnet werden[1]: Man unterstellt für die Kovarianzmatrix der Fehler F'_j (j=t-T,...,t), daß

$$\begin{pmatrix} E(F'^2_{t-T}) & E(F'_{t-T}F'_{t-T+1}) & \dots & E(F'_{t-T}F'_t) \\ E(F'_{t-T+1}F'_{t-T}) & E(F'^2_{t-T+1}) & \dots & E(F'_{t-T+1}F'_t) \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ E(F'_tF'_{t-T}) & E(F'_tF'_{t-T+1}) & \dots & E(F'^2_t) \end{pmatrix}$$

---

1 Vgl. etwa Johnston, a.a.O., S. 122 ff.

$$= \begin{pmatrix} \sigma^2 & 0 & \dots & 0 \\ 0 & \sigma^2 & \dots & 0 \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ 0 & 0 & \dots & \sigma^2 \end{pmatrix} , \tag{7}$$

daß also die Fehler konstante Varianzen besitzen und nicht miteinander korrelieren. Die Varianz σ^2 läßt sich unter diesen Annahmen als

$$\hat{\sigma}^2 = \frac{\mathbf{r'r}}{(T+1)-(k+2)} \tag{8}$$

schätzen[1] und für die Kovarianzmatrix der Schätzvariablen $\hat{W}_c, \hat{W}_0, \dots, \hat{W}_k$, nämlich

$$V(\hat{\mathbf{w}}) = \begin{pmatrix} V(\hat{W}_c) & C(\hat{W}_c\hat{W}_0) & \dots & C(\hat{W}_c\hat{W}_k) \\ C(\hat{W}_0\hat{W}_c) & V(\hat{W}_0) & \dots & C(\hat{W}_0\hat{W}_k) \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ C(\hat{W}_k\hat{W}_c) & C(\hat{W}_k\hat{W}_0) & \dots & V(\hat{W}_k) \end{pmatrix} \tag{9}$$

erhält man bei Gültigkeit von (7) den Schätzwert[2]

$$\hat{V}(\hat{\mathbf{w}}) = \sigma^2 \, (\mathbf{U'U})^{-1} . \tag{10}$$

Die gesuchten Varianzen $\hat{V}(\hat{W}_i)$ $(i=c,0,\dots,k)$ findet man auf der Hauptdiagonalen der Matrix $\hat{V}(\hat{\mathbf{w}})$.

Nun wurde bereits gezeigt, daß die Annahme der Homoskedastizität und der Unkorreliertheit nicht zutrifft. Die Fehler weisen vielmehr in jeder Periode individuell verschiedene Varianzen

---

1 Vgl. J. Johnston, a.a.O., S. 128.

2 Ebenda, S. 126.

auf und sind negativ autokorreliert. Man verschätzt sich systematisch, wenn man die Varianzen $\hat{V}(\hat{W}_i)$ nach (10) in Verbindung mit (8) berechnet und erhält damit falsche Konfidenzintervalle für die Parameter W_i. "Über Ausmaß und Richtung dieses Fehlers lassen sich kaum allgemeine Aussagen machen"[1].

Für den Spezialfall der linearen Einfachregression

$$Y = \alpha + \beta X + U$$

mit konstanter Varianz V(U) läßt sich zwar zeigen, daß bei negativer Autokorrelation der Störgrößen die Varianz der Regressionskoeffizienten gewöhnlich überschätzt wird, wenn die Berechnung gemäß (10) durchgeführt wird[2], ob aber auch im vorliegenden Fall eine Überschätzung von $\hat{V}(\mathbf{w})$ erfolgt, läßt sich nicht allgemein angeben. Überhaupt ist zur Zeit noch wenig darüber bekannt, welche Auswirkungen sich aus mehreren zugleich eintretenden Abweichungen vom klassischen Modell der MRA ergeben[3].

Daß Versuchen zur Ausschaltung der störenden Einflüsse der Heteroskedastizität und der Autokorrelation hier keine großen Erfolgsaussichten eingeräumt werden können, wird deutlich, wenn man die verallgemeinerte Methode der kleinsten Quadrate nach AITKEN betrachtet[4].

Unter Verwendung der Kovarianzmatrix

$$\mathbf{V} = \begin{pmatrix} E(F'^2_{t-T}) & E(F'_{t-T}F'_{t-T+1}) & \dots & E(F'_{t-T}F'_t) \\ E(F'_{t-T+1}F'_{t-T}) & E(F'^2_{t-T+1}) & \dots & E(F'_{t-T+1}F'_t) \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ E(F'_tF'_{t-T}) & E(F'_tF'_{t-T+1} & \dots & E(F'^2_t) \end{pmatrix} \quad (11)$$

1 H. Schneeweiß, a.a.O., S. 178.

2 Vgl. H. Gollnick: Einführung in die Ökonometrie, Stuttgart 1968, S. 70.

3 Vgl. H. Schneeweiß, a.a.O., S. 194.

4 Ebenda, S. 177 ff.

lassen sich die Regressionsparameter optimal als

$$\hat{w}_A = (U'V^{-1}U)^{-1}U'V^{-1}e \qquad (12)$$

schätzen, und die Varianzen $V(\hat{W}_{iA})$ findet man in der Matrix

$$V_A(\hat{w}_A) = (U'V^{-1}U)^{-1} \; . \qquad (13)$$

Da die Kovarianzmatrix V unbekannt ist, kann die verallgemeinerte Methode nur angewendet werden, wenn es gelingt, V zu schätzen, und die in der Literatur vorgeschlagenen Verfahren zur Berücksichtigung der Heteroskedastizität[1] und der Autokorrelation[2] der Fehler laufen auf diese Schätzung hinaus. Den Verfahren ist gemeinsam, daß sie bestimmte vereinfachende Annahmen über die Struktur der Matrix V treffen (etwa daß nur Heteroskedastizität oder nur Autokorrelation vorliegt), welche hier nicht aufrechterhalten werden können. Eine Schätzung von V ohne Zusatzkenntnisse erscheint im vorliegenden Fall nicht durchführbar. Damit kann auch nur die "normale" Schätzfunktion

$$\hat{w} = (U'U)^{-1}U'e$$

verwendet werden, und es besteht nicht die Möglichkeit, richtige Konfidenzintervalle für die Parameter W_c, W_0, ..., W_k zu bilden.

Aus demselben Grund kann nur eine Punkt- und nicht zusätzlich auch eine Intervallprognose für die Einnahmen einer bestimmten Periode abgegeben werden. Die Schätzung der Fehlervarianz nach

---

1 Vgl. etwa R. E. Park: Estimation with Heteroskedastic Error Terms, Econometrica 34 (1966), S. 888; H. C. Rutemiller u. D. A. Bowers: Estimation in a Heteroskedastic Regression Model, Journal of the American Statistical Association 63 (1968), S. 552 ff.

2 Vgl. etwa D. Cochrane u. G. H. Orcutt: Application of Least Squares Regression to Relationships Containing Autocorrelated Error Terms, Journal of the American Statistical Association 44 (1949), S. 32 ff; M. Grassini: The Transformation Matrix in the Linear Statistical Model with Serially Correlated Errors generated by Autoregressive Schemes, Statistica 33 (1973), S. 395 ff.

$$\sigma^2 = \frac{\mathbf{r}'\mathbf{r}}{(T+1) - (k+2)}$$

führt zu einer Mittelung der beobachteten quadratischen Abweichungen und erlaubt keine individuelle Betrachtung der Fehlervarianzen mehr.

Festzuhalten ist: Mit der MRA erhält man bei großer Anzahl von Rechnungserstellungen pro Periode praktisch erwartungstreue Schätzwerte für die Komponenten der Verweilzeitverteilung. Intervallschätzungen für die Parameter sind jedoch nicht möglich, da ohne Zusatzkenntnisse die zugehörigen Varianzen nicht fehlerfrei geschätzt werden können.

6.4 Beeinträchtigung der Schätzung durch Multikollinearität

Ein weiteres, technisches[1] Problem ist das der Multikollinearität. Die Schätzwerte für die Regressionsparameter ergaben sich als

$$\hat{\mathbf{w}} = (\mathbf{U}'\mathbf{U})^{-1}\mathbf{U}'\mathbf{e} \quad ,$$

wobei

$$\mathbf{U} = \begin{pmatrix} 1 & u_{t-T} & u_{t-T-1} & u_{t-T-2} & \cdots & u_{t-T-k} \\ 1 & u_{t-T+1} & u_{t-T} & u_{t-T-1} & \cdots & u_{t-T-k+1} \\ \cdot & \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & \cdot & & \cdot \\ 1 & u_t & u_{t-1} & u_{t-2} & \cdots & u_{t-k} \end{pmatrix} . \qquad (1)$$

Die Inverse von $\mathbf{U}'\mathbf{U}$ kann nur berechnet werden, wenn $\mathbf{U}$ den Rang k+2 besitzt. Ist eine Spalte der Matrix $\mathbf{U}$ von einer oder mehreren anderen linear abhängig, so ist $\mathbf{U}$ singulär und $\mathbf{U}'\mathbf{U}$ nicht invertierbar. Es liegt dann perfekte Multikollinearität

---

1 Vgl. D. E. Farrar u. R. R. Glauber: Multicollinearity in Regression Analysis: The Problem Revisited, Review of Economics and Statistics 49 (1967), S. 92 ff.

vor, und die Regressionskoeffizienten können nicht geschätzt werden.

Da die Umsatzmatrix **U** von der ersten Spalte abgesehen spaltenweise zusammenhängende Teile der Umsatzreihe $u_{t-T-k},\ldots,u_t$ enthält, hängt das Ausmaß der Multikollinearität von der Struktur dieser Reihe ab. Man sieht, daß perfekte Multikollinearität vorliegt, wenn k+1 aufeinanderfolgende Periodenumsätze stets in der festen Beziehung

$$u_j = a + b_1 u_{j-1} + b_2 u_{j-2} + \ldots + b_k u_{j-k}, \quad j=t-T,\ldots,t \tag{2}$$

stehen, d.h. wenn die zweite Spalte in **U** durch eine Linearkombination aus den übrigen erzeugt werden kann. Wegen der besonderen Struktur von **U** ist dann auch jede andere Spalte als Linearkombination der restlichen Spalten herstellbar.

Bezeichnet man die erklärenden Variablen, deren Ausprägungen die Matrix **U** spaltenweise enthält, mit U_c, U_0, $U_{-1},\ldots,U_{-k}$, so läßt sich der Fall der linearen Abhängigkeit der Spalten auch dadurch ausdrücken, daß die multiplen Korrelationskoeffizienten

$$R_{U_c \cdot U_0 U_{-1} \ldots U_{-k}},\ R_{U_0 \cdot U_c U_{-1} \ldots U_{-k}},\ \ldots,\ R_{U_{-k} \cdot U_c U_{-1} \ldots U_{-k+1}}$$

sämtlich den Wert Eins annehmen[1].

In der Praxis wird perfekte Multikollinearität in der Matrix **U** so gut wie nie zu beobachten sein. Eine Umsatzentwicklung, die exakt die Gleichung (2) erfüllt, ist nur schwer vorstellbar, und sobald eine einzige Abweichung von dieser Gleichung vor-

---

1 Vgl. D. E. Farrar u. R. R. Glauber, a.a.O., S. 98.

liegt, ist $\mathbf{U}$ nicht mehr singulär. $\mathbf{U'U}$ ist dann invertierbar, und man kann $\hat{\mathbf{w}}$ berechnen. Damit ist das Problem jedoch nur verlagert worden. Es stellt sich nun die Frage: Ist das Ausmaß der Multikollinearität hoch oder gering?

Hohe Multikollinearität tritt auf, wenn sich in der Umsatzmatrix $\mathbf{U}$ eine beinahe exakte lineare Abhängigkeit einer oder mehrerer erklärender Variablen von den übrigen ergibt. Sie hat zur Folge, daß man mit großen Schätzfehlern für die entsprechenden Regressionskoeffizienten rechnen muß[1]. In Einzelfällen wird $\hat{\mathbf{w}}$ keine Ähnlichkeit mit einer Verweilzeitverteilung mehr aufweisen. Die Schätzwerte $\hat{w}_i$ können aufgrund der Multikollinearität völlig unerwartete Werte annehmen und durchaus außerhalb des Intervalls von Null bis Eins liegen[2]. Eine Interpretation der Regressionskoeffizienten als Komponenten einer Verweilzeitverteilung ist dann nicht mehr möglich, und die Ergebnisse können nicht zur Analyse des Zahlungsverhaltens der Kunden herangezogen werden. Eine Prognose der zukünftigen Umsatzeinnahmen ist jedoch weiterhin unter einer bestimmten Bedingung möglich:

Stehen die zur Abgabe der Prognose $\hat{e}_t$ verwendeten Ausprägungen der erklärenden Variablen in etwa in derjenigen linearen Beziehung, die bei der Ermittlung der Regressionsparameter vorlag, so heben sich durch die Bildung der Linearkombination

$$\hat{e}_t = \hat{w}_c + \hat{w}_0 u_t + \hat{w}_1 u_{t-1} + \ldots + \hat{w}_k u_{t-k}$$

die vergrößerten Schätzfehler wieder auf, und der Prognosefehler kann auch bei hoher Multikollinearität klein ausfallen[3]. Bei einer starken Veränderung des linearen Zusammenhangs im Zeitablauf ist jedoch auch keine befriedigende Prognose mehr möglich[4].

---

1 Vgl. H. Schneeweiß, a.a.O., S. 136 f.
2 Vgl. J. C. Chambers, S. K. Mullick u. D. D. Smith, a.a.O.,S. 206.
3 Vgl. H. Schneeweiß, a.a.O., S. 137.
4 Ebenda.

Wann liegt nun ein schädlich hohes Ausmaß an Multikollinearität vor? Ein exaktes allgemeines Maß existiert nicht, wohl aber einige Daumenregeln[1]. Danach deuten die folgenden Anzeichen auf starke Multikollinearität hin:

1. Unerwartete (hier: negative) Vorzeichen der Regressionskoeffizienten.
2. Hohe paarweise Korrelation $r_{U_0U_{-1}}$, $r_{U_0U_{-2}}$, ..., $r_{U_{-k+1}U_{-k}}$ zwischen den erklärenden Variablen. Sämtliche Produktmomentkorrelationskoeffizienten sollten kleiner als 0,8 oder 0,9 sein.
3. Höhere Korrelation zwischen zwei erklärenden Variablen als die multiple Korrelation $R_{E.U_cU_0\ldots U_{-k}}$ zwischen der erklärten und den erklärenden Variablen.
4. Höhere multiple Korrelation zwischen den erklärenden Variablen als zwischen der erklärten und den erklärenden Variablen, d.h. wenn mindestens einer der Koeffizienten

$$R_{U_c.U_0U_{-1}\ldots U_{-k}},\ R_{U_0.U_cU_{-1}\ldots U_{-k}},\ \ldots,\ R_{U_{-k}.U_cU_0\ldots U_{-k+1}}$$

größer ist als $R_{E.U_cU_0\ldots U_{-k}}$.

Es reicht an sich aus, nur die letzte Regel anzuwenden. Sie deckt auch Fälle starker Multikollinearität auf, die mit den anderen Regeln nicht feststellbar sind. Dennoch haben auch die übrigen Vorschläge ihre Berechtigung. Stellt man etwa negative Vorzeichen bei der Berechnung von $\hat{w}$ fest oder starke paarweise Korrelation zwischen den erklärenden Variablen, so ist es gar nicht mehr erforderlich, den rechnerisch aufwendigen vierten Weg zu gehen, der nur dann nicht zu vermeiden ist, wenn die einfacheren Faustregeln keine hohe Multikollinearität erkennen lassen.

---

1 Vgl. D. E. Farrar u. R. R. Glauber, a.a.O., S. 98 f.

Bestimmte Formen der zur Schätzung $\hat{w}$ verwendeten Umsatzzeitreihe begünstigen von vornherein die Entstehung starker linearer Beziehungen zwischen den erklärenden Variablen. Weiter oben wurde darauf hingewiesen, daß perfekte Multikollinearität vorliegt, wenn die Zeitreihe die Gleichung

$$u_j = a + b_1 u_{j-1} + b_2 u_{j-2} + \ldots + b_k u_{j-k}$$

$$j = t\text{-}T,\ldots,t \qquad (3)$$

erfüllt. Dazu gehören auch die interessanten Spezialfälle

$$u_j = a \qquad j = t\text{-}T,\ldots,t, \qquad (4)$$

$$u_j = a + u_{j-1} \qquad j = t\text{-}T,\ldots,t, \qquad (5)$$

$$u_j = u_{j-d} \qquad j = t\text{-}T,\ldots,t;\ 1 \leq d \leq k, \qquad (6)$$

und $u_j = a + u_{j-d}$ $\qquad j = t\text{-}T,\ldots,t;\ 1 \leq d \leq k, \qquad (7)$

für die in Abbildung 19 jeweils ein Beispiel wiedergegeben ist.

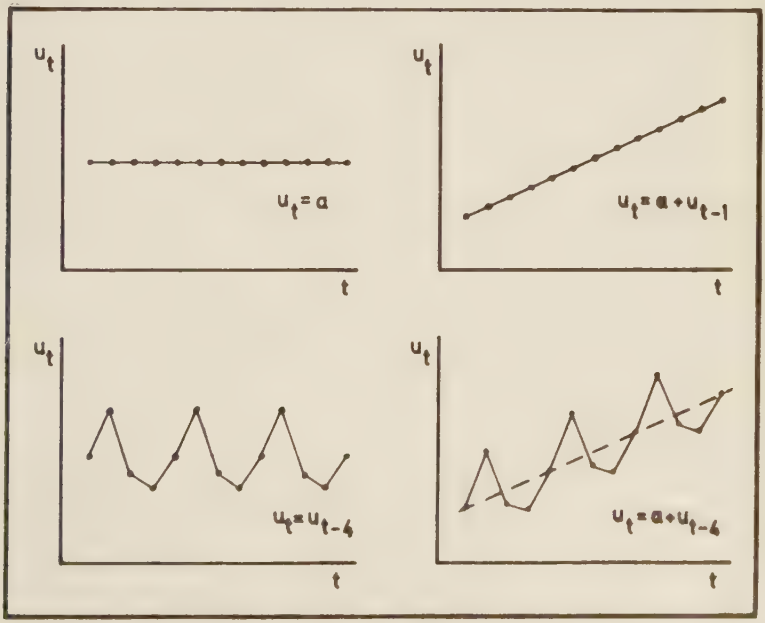

Abb. 19: Zeitreihen, die zu perfekter Multikollinearität führen

Eine exakte Übereinstimmung der tatsächlichen Umsatzreihe mit einer der idealtypischen Reihen der Abbildung ist zwar nur schwer vorstellbar, eine sehr enge Annäherung kann dagegen durchaus auftreten. In bestimmten Fällen läßt sich somit schon durch reine Betrachtung der verwendeten Umsatzzeitreihe erkennen, ob bei der Berechnung der Parameter mit einem größeren Ausmaß an Multikollinearität zu rechnen ist.

In diesem Zusammenhang kann es sich auch ausnahmsweise einmal nachteilig auf die Schätzgenauigkeit auswirken, wenn die Länge der Basisperiode zu groß gewählt wird; denn durch Vergrößerung des zugrundegelegten Intervalls kann bereits eine gewisse Glättung der Zeitreihe und damit eine Annäherung an den idealtypischen Verlauf $u_j = a$ bzw. $u_j = a + u_{j-1}$ bewirkt werden[1].

Was kann man nun tun, wenn tatsächlich unangenehm hohe Multikollinearität bei der Berechnung von $\hat{\mathbf{w}}$ auftritt? In der Literatur werden verschiedene Möglichkeiten zur Überwindung dieses Problems angeboten[2], auf die jedoch nicht eingeangen werden soll, weil sie der besonderen Struktur der Matrix $\mathbf{U}$ nicht Rechnung tragen. Gegen die Multikollinearität ist man hier machtlos[3]. Es bleibt nur die Hoffnung, daß die zur Einnahmeprognose verwendeten Periodenumsätze, die ja einen zusammenhängenden Teil der Umsatzreihe bilden, in ungefähr derselben linearen Beziehung stehen wie die Spalten der Umsatzmatrix $\mathbf{U}$, damit wenigstens die Prognosegenauigkeit nicht geschmälert wird.

---

1 Vgl. A. Nowack: Prognose bei unregelmäßigem Bedarf, in: P. Mertens (Hrsg.): Prognoserechnung, 2., neubearbeitete und erweiterte Auflage, Würzburg/Wien 1975, S. 85.

2 Vgl. z. B. die Zusammenstellung bei H. Schneeweiß, a.a.O.,S. 139 ff.

3 Eine Überwindung der Multikollinearität ist hier nur durch zusätzliche Information möglich. Wenn man z.B. weiß, wie sich die einzelnen Periodeneinnahmen e_t bezüglich der Verweilzeiten zusammensetzen, so kann man den multiplen Ansatz in k Einfachregressionen aufspalten und jede Komponente W_i getrennt schätzen. In diesem Fall könnte man allerdings auch das in Abschnitt 3.7.2 beschriebene Schätzverfahren verwenden. Der besondere Vorzug der MRA, nur die Zeitreihen der Umsätze und Einnahmen zur Schätzung der Verweilzeitverteilung zu benötigen, entfiele damit jedenfalls.

6.5 Bestimmung der Anzahl der in den Ansatz aufzunehmenden Variablen

Aus didaktischen Gründen wurde bislang davon ausgegangen, daß die Anzahl der verwendeten erklärenden Variablen bereits vor der Durchführung der Regressionsanalyse bekannt ist. Sofern man nicht über Zusatzkenntnisse verfügt, trifft das nicht zu. Vielmehr kann erst im Verlauf der eigentlichen Berechnung festgestellt werden, welche Variablen am Ende wirklich im Regressionsansatz enthalten sein sollen.

Es sei noch einmal an die Ausgangsüberlegung dieses Kapitels erinnert. Aufgrund der Erfahrung, daß Verweilzeitverteilungen am Ende häufig viele kleine Komponenten enthalten, die bei der Ermittlung mit der MRA einen großen Rechenaufwand verursachen, jedoch nur einen geringen Beitrag zur Erklärung der Einnahmen leisten, wird der Ansatz

$$e_t = W_0 u_t + W_1 u_{t-1} + \ldots + W_k u_{t-k} + W_{k+1} u_{t-k-1} + \ldots + W_n u_{t-n} + f_t \qquad (1)$$

umgeformt zu

$$e_t = W_c + W_0 u_t + W_1 u_{t-1} + \ldots + W_k r_{t-k} + f'_t \; . \qquad (2)$$

Darin wird W_c als derjenige Betrag interpretiert, welchen der "abgeschnittene" Teil $W_{k+1} u_{t-k-1} + \ldots + W_n u_{t-n}$ im Mittel annimmt, wobei k so gewählt werden soll, daß W_c einen gewünschten Wert nicht überschreitet. Die Wahl dieser Obergrenze von W_c entscheidet damit über die Anzahl der im Regressionsansatz der Grundgesamtheit enthaltenen erklärenden Variablen[1].

Entsprechend der Definition von W_c bietet sich folgendes Vorgehen zur Ermittlung von k an:

---

1 Die Richtung dieser Abhängigkeit ist üblicherweise entgegengesetzt und hier nur durch die besondere Situation zu begründen.

Man gibt einen Wert W_c^+ vor, etwa in Form eines als vernachlässigbar klein angesehenen absoluten Betrages oder eines Betrages, der nur einen geringen Prozentsatz der durchschnittlichen Periodeneinnahmen aus Umsätzen ausmacht. Dann schätzt man W_c aus dem Ansatz

$$e_t = W_c + W_0 u_t + f'_t \qquad (3)$$

und vergleicht $\hat{W}_c$ mit W_c^+. Ist $\hat{W}_c$ größer als W_c^+, so nimmt man $W_1 u_{t-1}$ in (3) auf, schätzt W_c aus dem Ansatz

$$e_t = W_c + W_0 u_t + W_1 u_{t-1} + f'_t \qquad (4)$$

und vergleicht erneut $\hat{W}_c$ mit W_c^+. Es werden nun solange neue Variablen aufgenommen, bis $\hat{W}_c$ unter W_c^+ gesunken ist und damit der endgültige Ansatz feststeht.

Da $\hat{W}_c$ ein Stichprobenwert ist und nicht der wahre Wert der Grundgesamtheit, ist diese Ermittlung von k theoretisch nicht exakt. Strenggenommen müßte man in jedem Schritt ein Konfidenzintervall für W_c ermitteln und die obere Grenze dieses Bereichs mit W_c^+ vergleichen. Aus schon genannten Gründen können aber keine Konfidenzintervalle für die Regressionsparameter bestimmt werden, so daß man mit dem etwas groberen Verfahren vorlieb nehmen muß.

Ein vergleichsweise ernster Fehler kann bei bestimmten Formen der Multikollinearität auftreten. Die Abbildung 20 zeigt die geometrische Darstellung einer Regression der Form

$$e_j = \hat{W}_c + \hat{W}_0 u_j + \hat{W}_1 u_{j-1} + r_j \ , \ j = t-T,\ldots,T \ , \qquad (5)$$

wobei zwischen den Ausprägungen der erklärenden Variablen eine starke Beziehung der Form $u_j = a + bu_{j-1}$ besteht.

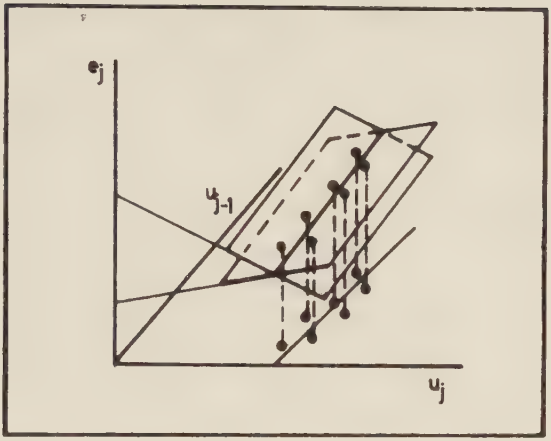

Abb. 20: Geometrische Darstellung einer Regression mit Multikollinearität, nach: H. Schneeweiß, a.a.O., S. 138.

Die Schätzung der Regression läßt sich geometrisch als Anpassung eines Punkthaufens durch eine Regressionsebene interpretieren[1]. Liegen die Beobachtungswerte in der (u_{j-1}, u_j) - Ebene dicht um die Gerade $u_j = a + bu_{j-1}$ geschart, so ist die Lage der Regressionsebene und damit auch ihr Schnittpunkt mit der e_j-Achse sehr stark vom Zufall abhängig. Der Achsenabschnitt $\hat{W}_c$ ist dann eine ausgesprochen ungenaue Schätzung des wahren Wertes W_c. Da trendartig steigende oder fallende Umsatzreihen zu ähnlichen Beziehungen zwischen den erklärenden Variablen führen, ist in solchen Fällen große Vorsicht bei der Anwendung des obigen Verfahrens zur Ermittlung von k geboten.

Nur wenn die lineare Beziehung der erklärenden Variablen durch den Koordinatenursprung verläuft, wird die Schätzung von W_c nicht durch Multikollinearität beeinflußt. Die Drehachse der Regres-

1 Vgl. H. Schneeweiß, a.a.O., S. 138.

sionsebene (vgl. Abb. 20) markiert dann nämlich selbst den Abschnitt $\hat{W}_c$ auf der Ordinate e_j. (Dieser Fall tritt z.B. bei zyklisch um ein konstantes Niveau schwankenden Umsätzen ein.) Sonst läßt sich bei starker Multikollinearität das absolute Glied nicht mehr inhaltlich interpretieren und kann nicht zur Bestimmung von k herangezogen werden.

Um dieses Problem zu umgehen, kann man anstelle des betragsmäßig kleinen Teils $W_{k+1},\ldots, W_n$ der Verweilzeitverteilung denjenigen heranziehen, welcher bei der Prognose den wesentlichen Beitrag zur Erklärung der Einnahmen beisteuert, also $W_0,\ldots,W_k$. Die Fragestellung lautet dann: Wieviele Variablen müssen in den Ansatz aufgenommen werden, damit eine befriedigende Erklärung der Einnahmen stattfindet? Zur Auswahl einer optimalen Regressionsgleichung werden in der Literatur verschiedene Verfahren angeboten[1]. Bis auf eines scheiden jedoch alle für eine Anwendung bei der Einnahmeprognose von vornherein aus, weil sie Signifikanztests verwenden, deren Voraussetzungen hier nicht gegeben sind[2].

Dieses Verfahren beurteilt die Regressionsansätze nach der Höhe des jeweiligen multiplen Bestimmtheitskoeffizienten, welcher den Anteil der durch die Regression erklärten Varianz an der Gesamtvarianz der abhängigen Variablen angibt[3]. Es lautet in für die vorliegende Fragestellung leicht modifizierter Form:

Man berechne die Regressionen der Ansätze

$$e_t = W_c + W_0 u_t + f'_t$$

$$e_t = W_c + W_0 u_t + W_1 u_{t-1} + f'_t$$

$$e_t = W_c + W_0 u_t + W_1 u_{t-1} + W_2 u_{t-2} + f'_t$$

---

1 Vgl. N. R. Draper u. H. Smith: Applied Regression Analysis, Third printing, New York/London/Sydney 1967, S. 163 ff.

2 Ebenda, S. 167 ff.

3 Ebenda, S. 164 ff.

usw.

und die zugehörigen Bestimmtheitskoeffizienten $R^2_{E.U_cU_o}$, $R^2_{E.U_cU_oU_{-1}}$, $R^2_{E.U_cU_oU_{-1}U_{-2}}$, usw. Mit jeder neu aufgenommenen Variablen wird sich ein höherer Koeffizient als der vorherige ergeben. Wegen der unterstellten Struktur der Verweilzeitverteilung - die letzten Komponenten werden als sehr klein angenommen - wird die Zunahme nach einigen Schritten nur noch sehr gering sein. Man bricht das Verfahren ab, wenn ein als genügend groß angesehener Prozentsatz der Varianz der Einnahmen erklärt wird, d.h. wenn $R^2_{E.U_cU_o\ldots U_{-k}}$ einen vorgegebenen Wert erreicht hat, oder wenn durch die Aufnahme weiterer unabhängiger Variablen praktisch kein Erklärungszuwachs mehr erzielt wird. Wann das der Fall ist, kann nur subjektiv entschieden werden; als Faustregel hält STRUCK einen Bestimmtheitskoeffizienten von mindestens 0,9 für erforderlich[1].

Da die Beurteilung des Regressionsansatzes mit Hilfe des Bestimmtheitsmaßes dasselbe schrittweise Vorgehen wie die nach Maßgabe des absoluten Gliedes $\hat{W}_c$ erfordert, kann man ohne großen Mehraufwand beide Kriterien zugleich heranziehen.

6.6 Zusammenfassende Beurteilung der Einsatzmöglichkeit

Die multiple Regressionsanalyse kann mit gewissen Einschränkungen zur Prognose der Einnahmen aus Umsätzen herangezogen werden. In günstigen Fällen stellt sie ein eigenständiges Verfahren zur Ermittlung der Verweilzeitverteilung dar, das jedoch keine Intervallschätzung der wahren Komponenten gestattet. Anders als bei den übrigen Verfahren zur Ermittlung des Liquidationsspektrums nimmt die Struktur der Umsatzzeitreihe Einfluß auf die Genauigkeit bei der Schätzung der Komponenten. Sie kann hohe Multikollinearität

---

1 Vgl. R. Struck: Kurzfristige statistische Vorausberechnung, Berlin 1973, S. 135.

verursachen und dadurch die Interpretation der Schätzwerte als Spektralkomponenten nicht mehr zulassen. Dann ist auch eine Prognose der Einnahmen nur noch in ganz bestimmten Fällen möglich. Hierin liegt eine deutliche Schwäche gegenüber der "normalen" Ermittlung der Verweilzeitverteilung.

Sofern nicht die Form der Umsatzreihe von vornherein auf die Entstehung starker Multikollinearität hindeutet, erscheint ein Versuch jedoch lohnend. Die MRA ist zwar sehr rechenaufwendig, kommt dafür aber ohne die arbeitsintensive Stichprobenentnahme aus und kann je nach der Datenorganisation und den vorhandenen technischen Hilfsmitteln (z.B. programmierbarem Rechner) bedeutend schneller durchführbar sein[1].

1 Vgl. R. Edin: Dynamische Analyse..., a.a.O., S. 180 f.

7 Schlußbemerkungen

Abschließend seien die wichtigsten Ergebnisse kurz zusammengefaßt. Mit der Verhältnisschätzung wurde ein Stichprobenverfahren gefunden, welches in der Durchführung einfacher als das NEUMEYERsche dreistufige Verfahren ist und praktisch unverzerrte Schätzungen für die Komponenten der Betragsverteilung liefert. Anhand eines Sonderfalles wurde ein Verfahren zur Ermittlung des Mindeststichprobenumfangs vorgestellt, welcher eine gewünschte Schätzgenauigkeit gewährleistet. Dabei wurde auch gezeigt, daß die Schätzung der Betragsverteilung bei gleichem Stichprobenumfang ungenauer ist als die der Anzahlverteilung und daß das Ausmaß des Unterschieds im wesentlichen vom Variationskoeffizienten der Rechnungsbeträge abhängt. Wenn auch die aufgezeigten funktionalen Zusammenhänge nur für den Fall der Unabhängigkeit zwischen Betrag und Verweilzeit der Rechnungen gelten, so ist doch zumindest größenordnungsmäßig eine entsprechende Aussage über den allgemeinen Fall möglich. Für die praktische Durchführung der Ermittlung wurden noch zwei Modifikationen vorgestellt, die auch die jüngeren Rechnungen der Stichprobe mitberücksichtigen können.

Im Rahmen der Betrachtung der Genauigkeit von Einnahmeprognosen mit der Verweilzeitverteilung wurde für den allgemeinen Fall gezeigt, daß die Prognose bei genügend großer Anzahl von Rechnungserstellungen pro Periode praktisch unverzerrt ist. Ausgehend vom Spezialfall der Unabhängigkeit von Betrag und Verweilzeit wurde demonstriert, daß die Verteilung des Prognosefehlers vielfach als annähernd normal angenommen werden kann und daß die Varianz des Fehlers als Maß für die Genauigkeit der Prognose unter anderem von der Anzahl der Rechnungserstellungen pro Periode, dem mittleren Rechnungsbetrag und der Varianz der Rechnungsbeträge abhängt. Dabei ist auch hier zu vermuten, daß im allgemeinen Fall ein ähnlicher Zusammenhang besteht. Ferner wurde ein praktikables Verfahren zur Abgabe von Intervallprognosen angegeben.

Aufbauend auf den Überlegungen zur Ermittlung der Verweilzeitverteilungen und der Genauigkeit der mit ihnen erstellten Prognosen wurden einige in der Literatur angebotene direkte und indirekte Stabilitätstests betrachtet. Es stellte sich heraus, daß ein exakter und zugleich praktikabler Test wegen der "Verklebung" der DM-Stücke zu Rechnungen nicht existiert. In Anlehnung an die Kontrollkartentechnik wurde aber ein Verfahren zur Stabilitätsüberwachung vorgestellt, das zumindest eine Entscheidungshilfe bei der Frage, ob eine Neuermittlung des Spektrums erfolgen soll oder nicht, bietet.

Schließlich wurde gezeigt, daß die multiple Regressionsanalyse eine praktisch unverzerrte Schätzung der Komponenten der Verweilzeitverteilung gestattet und in gewissen Fällen als eigenständiges Verfahren der Ermittlung angesehen werden kann. Sie stellt jedoch erheblich höhere Anforderungen an die formale Struktur der Inputdaten als die anderen Schätzverfahren und ist daher nicht immer anwendbar. Wenn die Anforderungen erfüllt sind, ist sie ausgesprochen komfortabel, da sie nur ein Minimum an Erhebungsaufwand erfordert.

Die Aussagen über die Verweilzeitverteilung erfolgten im wesentlichen unter den Annahmen, daß

1. viele Rechnungen pro Periode erstellt werden,
2. die Rechnungen voneinander unabhängig bezahlt werden,
3. die Verweilzeitverteilung stabil ist,
4. die in einer Periode erstellten Rechnungen bezüglich der Ausprägungen der Variablen R (Rechnungsbetrag) und B_i (nach i Perioden eingehender Betrag aus einer Rechnung) als Stichprobe aufgefaßt werden können.

Von besonderer Bedeutung ist die erste Annahme. Die Qualität der mit Hilfe von Verweilzeitverteilungen erstellten Prognosen hängt in hohem Maße von der Anzahl der Rechnungserstellungen pro Periode ab. Durch geeignete Wahl der Basisperiode werden sich viel-

fach in diesem Sinne günstige Voraussetzungen für die Anwendung der Prognose mit dem Liquidationsspektrum schaffen lassen. Die Basisperiode kann jedoch nicht beliebig groß gewählt werden, da sich sonst zu viele Zahlungsvorgänge innerhalb der als ohne zeitliche Ausdehnung behandelten Periode ereignen. Als Obergrenze kann wohl der Monat angesehen werden.

Die Annahme der Unabhängigkeit der Bezahlung verschiedener Rechnungen stellt auf die Kundenstruktur ab. Ihre Wirklichkeitsnähe kann nur im konkreten Einzelfall geprüft werden. Eine Untersuchung über die Folgen der Abhängigkeit bei der Bezahlung steht noch aus.

In der Stabilität der Verweilzeitverteilung äußert sich gleichbleibendes Zahlungsverhalten der Kunden. Verfügt das Unternehmen über einen gleichbleibenden Kundenstamm[1] und werden viele Rechnungen pro Periode erstellt, so ist diese Annahme plausibel. Bei ständig wechselnder Struktur der Kunden darf man gleichbleibendes Zahlungsverhalten im allgemeinen nicht erwarten; da es sich bei der Einnahmeprognose mittels Verweilzeitverteilungen um ein Verfahren der kurzfristigen Prognose handelt, kann man aber davon ausgehen, daß sich die Veränderung der Struktur des Kundenkreises in der interessierenden Zeitspanne eher evolutorisch vollzieht und somit auch die Verweilzeitverteilung sich nur allmählich verändert. Die Annahme der Stabilität ist dann zumindest annähernd erfüllt. Dazu kommt, daß durch eigene Maßnahmen, etwa durch Veränderung der Zahlungsbedingungen oder Verschärfung des Mahnwesens, unliebsamen Entwicklungen des Zahlungsverhaltens der Kunden in gewissem Umfang entgegengewirkt werden kann[2]. Da sich die Prognose mit dem Liquidationsspektrum ohnehin recht ro-

---

1 Vgl. H. Kossbiel, a.a.O., S. 86 ff.

2 Vgl. L. Engels, a.a.O., S. 134 ff.

bust in bezug auf Abweichungen des verwendeten Spektrums vom wahren Spektrum erweist[1], ist diese Prämisse aus der Sicht der praktischen Erfordernisse meistens erfüllt.

Die vierte Annahme wird für die Aussagen über die Eigenschaften des Prognosefehlers benötigt. Genau wie die vorherige Prämisse erscheint sie plausibel bei gleichbleibendem Kundenstamm und großer Anzahl von Rechnungserstellungen pro Periode. Etwas grob kann man sie bei Gültigkeit der dritten Voraussetzung auch als annähernde Konstanz von Mittelwert und Varianz der Rechnungsbeträge der einzelnen Perioden bezeichnen. Auch wenn die tatsächlichen Ausprägungen dieser Parameter sich über die Perioden hinweg verändern, so ist doch ein gewisser Ausgleich dadurch gegeben, daß in den Prognosefehler und in seine Varianz immer die Werte mehrerer Perioden eingehen[2]. Auch hier braucht aus der Sicht der praktischen Anwendung somit nicht unbedingt eine exakte Erfüllung der Annahme gegeben zu sein. Wie groß der Spielraum tatsächlich ist, bleibt allerdings offen.

---

1 Vgl. S. 100.

2 Vgl. Formel (12) auf S. 102.

Literaturverzeichnis

BATSCHELET, E.	Über eine Kontingenztafel mit fehlenden Daten, Biometrische Zeitschrift 2 (1960), S. 236 - 243.
BENISHAY, Haskel	Managerial Controls of Accounts Receivable: A Deterministic Approach, Journal of Accounting Research 3 (1965), S. 114 - 132.
Derselbe	Managerial Controls of Accounts Receivable: A Reply, Journal of Accounting Research 4 (1966), S. 131 - 132.
Derselbe	A Stochastic Model of Credit Sales Debt, Journal of the American Statistical Association 61 (1966), S. 1010 - 1028.
Derselbe	Parameters and Relations of Stochastically Lagged and Disaggregative Time Series, Econometrica 36 (1968), S.155 - 171.
BEYER, William H. (ed.)	CRC Handbook of Tables for Probability and Statistics, 2nd edition, Cleveland 1968.
BROWN, Robert Goodell	Smoothing, Forecasting and Prediction of Discrete Time Series, Englewood Cliffs., N. J. 1963.
CHAMBERS, John C., MULLICK, Satinder K. und SMITH, Donald D.	An Executive Guide to Forecasting, New York 1974.
CLAUSS, Günter und EBNER, Heinz	Grundlagen der Statistik für Psychologen, Pädagogen und Soziologen, 2., neubearbeitete und erweiterte Auflage, Zürich/Frankfurt a. M. 1975.
COCHRAN, William G.	Stichprobenverfahren, Berlin/New York 1972.
COCHRANE, D. und ORCUTT, G. H.	Application of Least Squares Regression to Relationships Containing Autocorrelated Error Terms, Journal of the American Statistical Association 44 (1949), S. 32 - 61.
DEMING, William Edwards	Some Theory of Sampling, New York/London 1950.

DRAPER, N. R. und SMITH, H.	Applied Regression Analysis, Third printing, New York/London/Sydney 1967.
EDIN, Robert	Wirkungsweise und Voraussetzungen der Prognose mittels Verweilzeitverteilungen - Ergebnisse einiger Experimente mit Zufallszahlen, Zeitschrift für Betriebswirtschaft 38 (1968), S. 743 - 764.
Derselbe	Übergangsfunktionen in betriebswirtschaftlichen Systemen, Zeitschrift für Betriebswirtschaft 39 (1969), S. 569 - 594.
Derselbe	Dynamische Analyse betrieblicher Systeme - Ein Beitrag zur industriellen Planung, Berlin 1971.
EDIN, Robert und SCHMITT, Hermann J.	Verweilzeitverteilungen und Prognosen - Einige empirische Ergebnisse, Zeitschrift für betriebswirtschaftliche Forschung 21 (1969), S. 484 - 506.
ENGELS, Ludwig	Zur Optimierung einer zeitabhängigen Liquiditätspolitik bei Saisonunternehmen, Dissertation Köln 1970.
ENRICK, N. L. und SCHÄFER, W.	Quantitative Marktprognose, Heidelberg 1972.
FARRAR, Donald E. und GLAUBER, Robert R.	Multicollinearity in Regression Analysis: The Problem Revisited, Review of Economics and Statistics 49 (1967), S. 92 - 107.
FRISCH, Ragnar	On the Notation of Equilibrium and Disequilibrium, The Review of Economic Studies 3 (1935/36), S. 100 - 105.
GAHSE, Sigfrid	Liquiditätsprognosen auf der Grundlage von Phasenfolgen (mit Hilfe von EDVA), Dissertation Mannheim 1967.
Derselbe	Dynamische Liquiditätsplanung DYPOL auf IBM System/360, IBM Form 8 064 200.
Derselbe	Eine Systematik der Vorhersageverfahren, IBM Form 81 564.

GEPPERT, M. P.	Erwartungstreue plausibelste Schätzer aus dreieckig gestutzten Kontingenztafeln, Biometrische Zeitschrift 3 (1961), S. 54 - 67.
GIESE, Olaf	Das betriebswirtschaftliche Problem der Liquidität, Dissertation Berlin 1971.
GOLDBERGER, Arthur S.	Econometric Theory, New York/London/ Sydney 1965.
GOLLNICK, Heinz	Einführung in die Ökonometrie, Stuttgart 1968.
GOOD, I. J.	What are Degrees of Freedom? The American Statistician 27 (1973), S. 227 - 228.
GRAF, Heinz	Die kurzfristige Absatzprognose als Bestandteil der Planung in der Konsumgüterindustrie, Dissertation Zürich 1968.
GRANT, Eugene L.	Statistical Quality Control, 3rd edition, New York 1964.
GRASSINI, Maurizio	The Transformation Matrix in the Linear Statistical Model with Serially Correlated Errors generated by Autoregressive Schemes, Statistica 33 (1973), S. 395 - 401.
GRÜNEWÄLDER, Otto	Mittel der Liquiditätskontrolle in industriellen Unternehmungen,Dissertation Mannheim 1965.
HANSEN, Morris H., HURWITZ, William N. und MADOW, William G.	Sample Survey Methods and Theory, Volume 1, Methods and Applications, New York/ London/Sydney 1953.
HAUSCHILDT, Jürgen	Organisation der finanziellen Unternehmensführung - Eine empirische Untersuchung, Stuttgart 1970.
HAUSTEIN, Heinz-Dieter	Prognoseverfahren in der sozialistischen Wirtschaft, Berlin 1970.
HAYS, William L.	Statistics for the Social Sciences, Second edition, London 1974.
HEWARD, James H. und STEELE, Peter M.	Business Control trough Multiple Regression Analysis - A Technique for the Numerate Manager, Haverhill 1973.
JOHNSTON, J.	Econometric Methods, 2nd edition, Tokyo 1972.

KAISER, Dietmar — Optimale Liquiditätsplanung mit Hilfe der elektronischen Datenverarbeitung und ihre Anwendung in Brauereiunternehmen, Dissertation München 1973.

KELLERER, Hans — Theorie und Technik des Stichprobenverfahrens - Eine Einführung unter besonderer Berücksichtigung der Anwendung auf soziale und wirtschaftliche Massenerscheinungen, Einzelschriften der Deutschen Statistischen Gesellschaft, 3. Auflage, München 163.

KISH, Leslie und HESS, Irene — On Variances of Ratios and their Differences in Multi-Stage Samples, Journal of the American Statistical Association 54 (1959), S. 416 - 446.

KMENTA, Jan — Elements of Econometrics, New York 1971.

KOCH, Helmut — Betriebliche Planung - Grundlagen und Grundfragen der Unternehmenspolitik, Wiesbaden 1961.

KOCKELKORN, Götz — Verweilzeitverteilungen und Prognosen im betrieblichen Produktionsbereich, Zeitschrift für betriebswirtschaftliche Forschung 13 (1971), S. 83 - 105.

KOEBEL, Dietrich — Finanzplanung in der Schuhindustrie, Zeitschrift für handelswissenschaftliche Forschung, Neue Folge 7 (1955), S. 291 - 293.

KOSSBIEL, Hugo — Die Umsatzeinnahmen als Gegenstand der unternehmerischen Liquiditätsplanung und Liquiditätspolitik, Berlin 1968.

KREYSZIG, Erwin — Statistische Methoden und ihre Anwendungen, 5., erweiterte Auflage, Göttingen 1975.

KUNKEL, Dieter — Grundlagen der betrieblichen Gesamtplanung (Ein Versuch zu einer Theorie der laufenden Gesamtplanung im Industriebetrieb), Dissertation Berlin 1960.

LANGE, Heinz — Die Vorausbestimmung des Absatzes als Grundlage der unternehmerischen Planung - Dargestellt am Beispiel einer Handelsauskunftei, Dissertation Erlangen-Nürnberg 1968.

LANGEN, Heinz	Die Prognose von Zahlungseingängen - Die Abhängigkeit der Bareinnahmen von Umsätzen und Auftragseingängen in dynamischer Betrachtung, Zeitschrift für Betriebswirtschaft 34 (1964), S. 289 - 326.
Derselbe	Betriebliche Zahlungsströme und ihre Planung in dynamischer Sicht, Zeitschrift für Betriebswirtschaft 35 (1965), S. 261 - 279.
Derselbe	Grundzüge einer betriebswirtschaftlichen Dispositionsrechnung, Zeitschrift für Betriebswirtschaft 36 (1966), Ergänzungsheft 1, S. 71 - 81.
Derselbe	Einige Werkzeuge betrieblicher Konjunkturtheorie, Zeitschrift für Betriebswirtschaft 37 (1967), S. 553 - 566.
Derselbe	Die betriebswirtschaftliche Dispositionsrechnung im Dienste der Unternehmensführung, in: Rudolf W. STÖHR (Hrsg.): Unternehmensführung auf neuen Wegen, Wiesbaden 1967, S. 231 - 243.
Derselbe	Preis- und konditionspolitische Maßnahmen in der Rezession, in: Herbert JACOB (Hrsg.): Schriften zur Unternehmensführung, Band 1: Unternehmenspolitik bei schwankender Konjunktur, Wiesbaden 1967, S. 43 - 60.
LANGEN, Heinz und WEINTHALER, Fritz	Prognose mit Hilfe von Verweilzeitverteilungen, in: Peter MERTENS (Hrsg.): Prognoserechnung, 2., neubearbeitete und erweiterte Auflage, Würzburg/Wien 1975, S. 101 - 118.
LANGEN, Heinz, EDIN, Robert, KOCKELKORN, Götz, SCHMITT, Hermann und WEINTHALER, Fritz	Unternehmensplanung mit Verweilzeitverteilungen - Eine Anleitung für Praktiker, Berlin 1971.
LEWANDOWSKI, Rudolf	Prognose- und Informationssysteme und ihre Anwendungen, Band 1, Berlin/New York 1974.
Derselbe	Prognosesysteme und ihre heutigen Anwendungsmöglichkeiten in der Unternehmensführung, in: G. BAMBERG und O. OPITZ: Information und Prognose, Meisenheim am Glan 1975, S. 49 - 89.

LIENERT, G. A.	Verteilungsfreie Methoden in der Biostatistik, Zweite völlig neu bearbeitete Auflage, Band I, Meisenheim am Glan 1973.
LUDWIG, Heinz	Die Einnahmeseite des Finanzplans - Eine Budgetstudie, Die Betriebswirtschaft 2 (1932), S. 33 - 36.
LÜCKE, Wolfgang	Finanzplanung und Finanzkontrolle, Wiesbaden 1962.
LÜTTICKEN, Frank H.	Praktische Erfahrungen mit dem Finanzplanungsmodell DYPOL, IBM-Nachrichten 19 (1969), S. 771 - 774.
MENGES, Günter und SKALA, Heinz J.	Grundriß der Statistik, Teil 2: Daten - Ihre Gewinnung und Verarbeitung, Opladen 1973.
MEYER, R. M.	On Correlation and Linearity, The American Statistician 28 (1974), S. 103.
MOOD, Alexander M., GRAYBILL, Franklin A. und BOES, Duane C.	Introduction to the Theory of Statistics, Third edition, Tokyo 1974.
NEITZEL, Willi	Finanzplanung in der Textilindustrie, Zeitschrift für handelswissenschaftliche Forschung, Neue Folge 7 (1955), S. 294 - 296.
NEUMEYER, Ludwig	Zum Problem der Ermittlung von Liquidationsspektren und ihrer Stabilität im Zeitablauf, Dissertation Mannheim 1966.
NOWACK, Arthur	Prognose bei unregelmäßigem Bedarf, in: Peter MERTENS (Hrsg.): Prognoserechnung, 2., neubearbeitete und erweiterte Auflage, Würzburg/Wien 1975.
OLMSTAED, P. S.	Distribution of Sample Arrangements for Runs up and down, The Annals of Mathematical Statistics 17 (1946), S. 24 - 33.
OPITZ, Lieselotte	Prognosen in der Marktforschung - Die Eignung der demoskopischen Marktforschung zur Formulierung von Prognosen, Wiesbaden 1969.
ORTH, Ludwig	Die kurzfristige Finanzplanung industrieller Unternehmungen, Köln/Opladen 1961.

PARK, R. E.	Estimation with Heteroskedastic Error Terms, Econometrica 34 (1966), S. 888.
PÖNNINGHAUS, Siegfried	Betriebswirtschaftliche Multiplikatoren, Zeitschrift für betriebswirtschaftliche Forschung 19 (1967), S. 659 - 673.
RECKSIEGEL, Wolf-Rüdiger	Die Anwendung der Regressions- und Korrelationsanalyse in der Kostenrechnung, Dissertation Münster 1972.
REIF, K.	Bedarfsvorhersage mittels mathematisch-statistischer Verfahren, IBM-Form 81 518.
ROGGE, Hans-Jürgen	Methoden und Modelle der Prognose aus absatzwirtschaftlicher Sicht - Ein Beitrag zur Prognoseforschung im Unternehmensbereich, Berlin 1972.
ROTHSCHILD, Kurt W.	Wirtschaftsprognose - Methoden und Probleme, Berlin/Heidelberg/New York 1969.
Derselbe	Zur Frage der Erfolgsbeurteilung ökonomischer Prognosen, Zeitschrift für die gesamte Staatswissenschaft 130 (1974), S. 577 - 586.
RUTEMILLER, Herbert C. und BOWERS, David A.	Estimation in a Heteroskedastic Regression Model, Journal of the American Statistical Association 63 (1968), S. 552 - 557.
SACHS, Lothar	Statistische Auswertungsmethoden, Zweite, neubearbeitete und erweiterte Auflage, Berlin/Heidelberg/New York 1969.
SCHAAFSMA, A. H. und WILLEMZE, F. G.	Moderne Qualitätskontrolle - Statistische und organisatorische Grundlagen der Qualitätsgestaltung, 4. Auflage, Eindhoven 1961.
SCHIEL, Hermann A.	Die Liquiditätssteuerung als kybernetisches System, Dissertation München 1972.
SCHMETTERER, Leopold	Einführung in die mathematische Statistik, Zweite, verbesserte und wesentlich erweiterte Auflage, Wien/New York 1966.

SCHMITT, Hermann J.	Die Planrevision in dynamischen Abgangsfolgemodellen, Betriebswirtschaftliche Forschung und Praxis 20 (1968), S. 24 - 37.
Derselbe	Planungsbuchhaltung - Ein Beitrag zur Integration von Buchhaltung und Planung, Berlin 1971.
SCHNEEWEISS, Hans	Ökonometrie, Würzburg/Wien 1971.
SCHRÖDER, Michael	Einführung in die kurzfristige Zeitreihenprognose und Vergleich der einzelnen Verfahren, in: Peter MERTENS (Hrsg.): Prognoserechnung, 2., neubearbeitete und erweiterte Auflage, Würzburg/Wien 1975.
SCHÜTZ, Waldemar	Methoden der mittel- und langfristigen Prognose - Eine Einführung, München 1975.
SCHWARZE, Jochen	Probleme der Fehlermessung bei quantitativen ökonomischen Prognosen, Zeitschrift für die gesamte Staatswissenschaft 129 (1973), S. 535 - 558.
SIEGEL, Sidney	Nonparametric Statistics for the Behavioral Sciences, Tokyo 1956.
SIMMONS, David A.	Practical Quality Control, Reading 1970.
SLAKTER, Malcolm J.	A Comparison of the Pearson Chi-Square and Kolmogorov-Goodness-Of-Fit Tests with Respect to Validity, Journal of the American Statistical Association 60 (1965), S. 854 - 858.
STATISTISCHES BUNDESAMT (Hrsg.)	Stichproben in der amtlichen Statistik, Stuttgart/Mainz 1960.
STRUCK, Richard	Kurzfristige statistische Vorausberechnung, Berlin 1973.
THEIL, Henry	Economic Forecasts and Policy, Second revised edition, Amsterdam 1965.
Derselbe	Applied Economic Forecasting, Amsterdam 1966.
THIEMANN, Klaus R.	Das Liquiditätsproblem in der leistungswirtschaftlich orientierten Planung unter besonderer Berücksichtigung der Auftragsfertigung, Dissertation Hamburg 1972.

TURNER, John	Forecasting Practices in British Industry, London 1974.
WESTERN ELECTRIC	Statistische Qualitätskontrolle - Das Handbuch der Western Electric Company für den Praktiker, Stuttgart 1969.
WILKS, Samuel S.	Mathematical Statistics, Second printing with corrections, New York/London 1963.
WITTE, Eberhard und KLEIN, Herbert	Finanzplanung der Unternehmung/Prognose und Disposition, Reinbek 1974
WONNACOTT, Thomas H. und WONNACOTT, Ronald J.	Introductory Statistics for Business and Economics, New York 1972.
WOODWARD, R. H. und GOLDSMITH, P. L.	Cumulative Sum Techniques, I. C. I. Monograph No. 3, Edinburgh 1964.
WRIGHT, F. K.	Managerial Controls of Accounts Receivable: A Comment, Journal of Accounting Research 4 (1966), S. 127 - 130.
YAMANE, Taro	Statistik - Ein einführendes Lehrbuch, Band 2, Frankfurt a. M. 1976.
YATES, Frank	Sampling Methods for Census and Surveys, Third edition, Second impression, London 1965.
ZAUNERT, Jörg	Die Liquidität der Betriebswirtschaft als Problem der Regelung und Steuerung des Zahlungs- und Kreditverkehrs, Dissertation Braunschweig 1975.

Lebenslauf

Geburtstag	23.11.1946
Geburtsort	Bielefeld
Staatsangehörigkeit	deutsch
Eltern	Reinhard Dietrich, Kaufmann und Elisabeth Dietrich, geb. Binsker
Schulen	1953 - 1957 Volksschule in Bielefeld 1957 - 1963 Gymnasium in Bielefeld 1963 - 1966 Gymnasium in Bad Salzuflen
Reifeprüfung	21.2.1966 in Bad Salzuflen
Studium	1.4.1968 - 31.10.1973 Betriebswirtschaftslehre an der Universität Münster
Diplomprüfung	19.10.1973
Berufstätigkeit	Seit 1.11.1973 Assistent am Institut für Ökonometrie und Wirtschaftsstatistik der Universität Münster